

# 腾讯云数据仓库 TCHouse-D

## 常见问题

## 产品文档



腾讯云

**【版权声明】**

©2013-2024 腾讯云版权所有

本文档著作权归腾讯云单独所有，未经腾讯云事先书面许可，任何主体不得以任何形式复制、修改、抄袭、传播全部或部分本文档内容。

**【商标声明】**

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。

**【服务声明】**

本文档意在向客户介绍腾讯云全部或部分产品、服务的当时的整体概况，部分产品、服务的内容可能有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

---

## 文档目录

常见问题

    常见操作问题

    常见报错

# 常见问题

## 常见操作问题

最近更新时间：2024-07-31 09:20:12

### 腾讯云数据仓库 TCHouse-D 是否支持修改列名？

在 1.2.0 版本之后, 开启 `"light_schema_change"="true"` 选项时, 可以支持修改列名。

在 1.2.0 版本之前或未开启 `"light_schema_change"="true"` 选项时, 不支持修改列名。

在腾讯云数据仓库 TCHouse-D 使用的 Doris 内核中, 列名称是直接写入到数据文件中的, 在查询时会直接通过列名查找到对应的列。因此, 改列名不仅是简单的元数据修改, 还会涉及到数据的重写, 是一个非常重的操作。开启 `light schema change` 后可支持修改列名, 但该功能目前尚不成熟, 建议谨慎使用。

### Unique Key 模型的表是否支持创建物化视图？

不支持。

Unique Key 模型的表是一个对业务比较友好的表, 因为其特有的按照主键去重的功能, 能够很方便的同步数据频繁变更的业务数据库。因此, 很多用户在将数据接入时会首先考虑使用 Unique Key 模型。

但遗憾的是, **Unique Key 模型的表是无法建立物化视图的**。原因在于, 物化视图的本质, 是通过预计算来将数据“预先算好”, 这样在查询时直接返回已经计算好的数据, 来加速查询。在物化视图中, “预计算”的数据通常是一些聚合指标, 例如求和、求 count。这时, 如果数据发生变更, 如 `update` 或 `delete`, 因为预计算的数据已经丢失了明细信息, 因此无法同步的进行更新。例如一个求和值5, 可能是 1+4, 也可能是2+3。因为明细信息的丢失, 我们无法区分这个求和值是如何计算出来的, 因此也就无法满足更新的需求。

### show backends/frontends 查看到的信息不完整？

在执行如 `show backends/frontends` 等某些语句后, 结果中可能会发现有部分列内容不全。例如 `show backends` 结果中看不到磁盘容量信息等。

通常这个问题会出现在集群有多个 FE 的情况下, 如果用户连接到非 Master FE 节点执行这些语句, 就会看到不完整的信息。这是因为此类信息仅存在于 Master FE 节点。例如 BE 的磁盘使用量信息等。所以只有在直连 Master FE 后, 才能获得完整信息。

当然, 用户也可以在执行这些语句前, 先执行 `set forward_to_master=true;` 这个会话变量设置为 true 后, 后续执行的一些信息查看类语句会自动转发到 Master FE 获取结果。这样, 不论用户连接的是哪个 FE, 都可以获取到完整结果了。

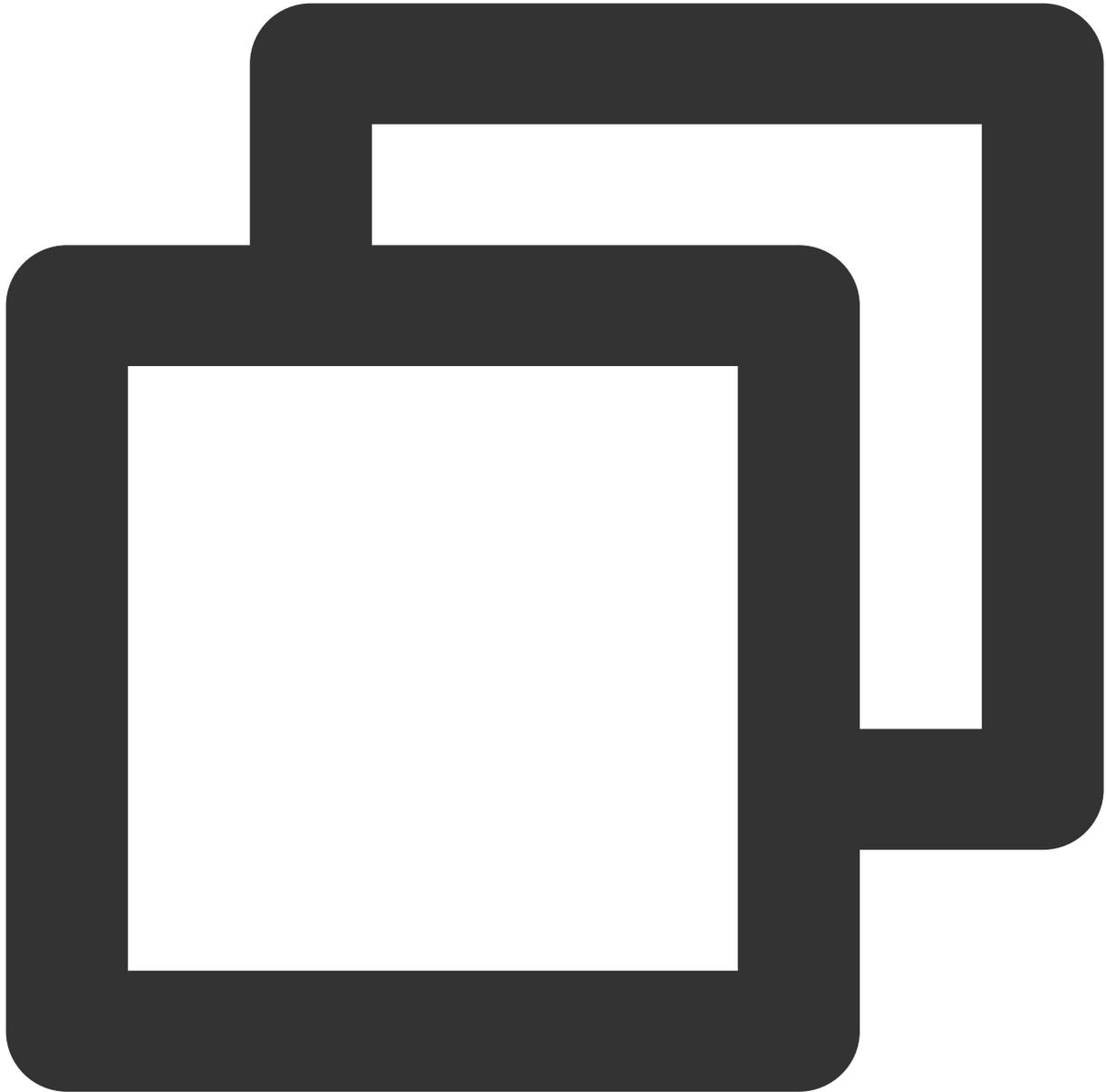
### Unique Key 模型查询结果不一致？

某些情况下, 当用户使用相同的 SQL 查询一个 Unique Key 模型的表时, 可能会出现多次查询结果不一致的现象。并且查询结果总在 2-3 种之间变化。

这可能是因为在同一批导入数据中, 出现了 key 相同但 value 不同的数据, 这会导致不同副本间因数据覆盖的先后顺

序不确定而产生的结果不一致的问题。

例如表定义为 k1, v1。一批次导入数据如下：



```
1, "abc"  
1, "def"
```

那么可能副本1的结果是 1, "abc"，而副本2的结果是 1, "def"。从而导致查询结果不一致。为了确保不同副本之间的数据先后顺序唯一，可以参考 [Sequence 列](#) 功能。

# 常见报错

最近更新时间：2024-07-31 09:20:25

## 查询报错：Failed to get scan range, no queryable replica found in tablet: xxxx

这种情况是因为对应的 tablet 没有找到可以查询的副本，通常原因可能是 BE 宕机、副本缺失等。可以先通过 `show tablet tablet_id` 语句，然后执行后面的 `show proc` 语句，查看这个 tablet 对应的副本信息，检查副本是否完整。同时还可以通过 `show proc "/cluster_balance"` 信息来查询集群内副本调度和修复的进度。

关于数据副本管理相关的命令，可以参阅 [数据副本管理](#)。

## tablet writer write failed, tablet\_id=27306172, txn\_id=28573520, err=-235 or -215 or -238

这个错误通常发生在数据导入操作中。错误码为 -235 或 -238。

-235 错误的含义是，对应 tablet 的数据版本超过了最大限制（默认500，由 BE 参数 `max_tablet_version_num` 控制），后续写入将被拒绝。例如问题中这个错误，即表示 27306172 这个 tablet 的数据版本超过了限制。通常是因为导入的频率过高，大于后台数据的 `Compaction` 速度，导致版本堆积并最终超过了限制。此时，可以先通过 `show tablet 27306172` 语句，然后执行结果中的 `show proc` 语句，查看 tablet 各个副本的情况。结果中的 `versionCount` 即表示版本数量。如果发现某个副本的版本数量过多，则需要降低导入频率或停止导入，并观察版本数是否有下降。如果停止导入后，版本数依然没有下降，则需要查看对应 BE 节点的 `be.INFO` 日志，搜索 `tablet id` 以及 `Compaction` 关键词，检查 `Compaction` 是否正常运行。

-238 错误通常出现在同一批导入数据量过大的情况，从而导致某一个 tablet 的 `Segment` 文件过多（默认 200，由 BE 参数 `max_segment_num_per_rowset` 控制）。此时建议减少一批次导入的数据量，或者适当提高 BE 配置参数值来解决。

## tablet 110309738 has few replicas: 1, alive backends: [10003]

这个错误可能发生在查询或者导入操作中。通常意味着对应 tablet 的副本出现了异常。

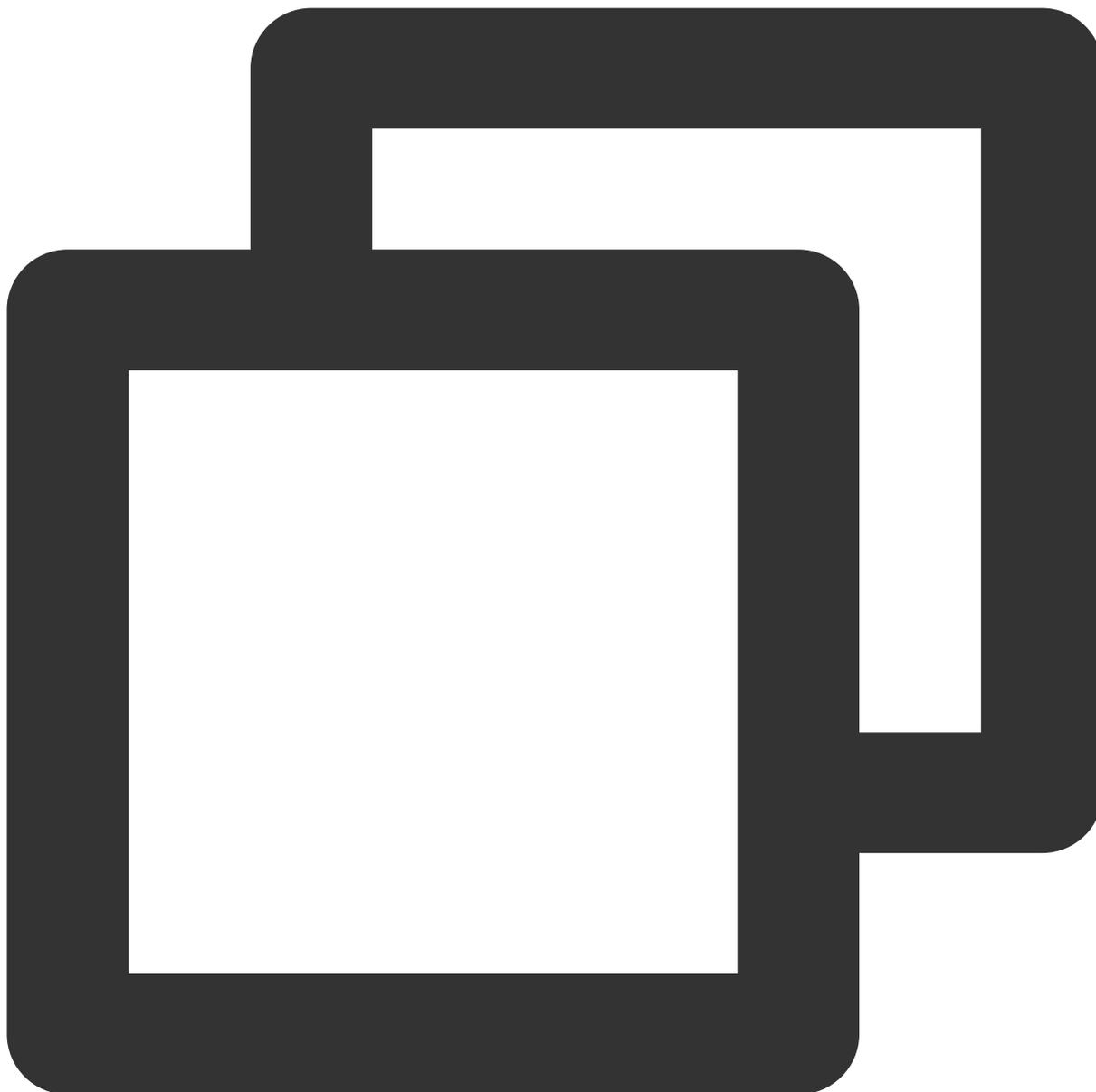
此时，可以先检查 BE 节点是否有宕机，如 `isAlive` 字段为 `false`，或者 `LastStartTime` 是最近的某个时间（表示最近重启过）。如果 BE 有宕机，则需要通过工单[联系我们](#)排查问题。如果没有 BE 节点宕机，则需要通过 `show tablet 110309738` 语句，然后执行结果中的 `show proc` 语句，查看 tablet 各个副本的情况，进一步排查。

## disk xxxxx on backend xxx exceed limit usage

通常出现在导入、Alter 等操作中。这个错误意味着对应 BE 的对应磁盘的使用量超过了阈值（默认95%）此时可以先通过 `show backends` 命令，其中 `MaxDiskUsedPct` 展示的是对应 BE 上，使用率最高的那块磁盘的使用率，如果超过95%，则会报这个错误。您可选择手动删除部分数据释放空间，或者扩容云盘解决此问题。若磁盘使用率异常增高，可通过工单[联系我们](#)具体排查。

## -214 错误

在执行导入、查询等操作时，可能会遇到如下错误：



```
failed to initialize storage reader. tablet=63416.1050661139.aa4d304e7a7aff9c-f0fa7
```

-214 错误意味着对应 tablet 的数据版本缺失。例如上述错误，表示 tablet 63416 在 192.168.100.10 这个 BE 上的副本的数据版本有缺失。（可能还有其他类似错误码，都可以用如下方式进行排查和修复）。

通常情况下，如果您的数据是多副本的，那么系统会自动修复这些有问题的副本。可以通过以下步骤进行排查：

1. 通过 `show tablet 63416` 语句并执行结果中的 `show proc xxx` 语句来查看对应 tablet 的各个副本情况。通常我们需要关心 `Version` 这一列的数据。

正常情况下，一个 tablet 的多个副本的 `Version` 应该是相同的。并且和对应分区的 `VisibleVersion` 版本相同。

- 您可以通过 `show partitions from tblx` 来查看对应的分区版本（`tablet` 对应的分区可以在 `show tablet` 语句中获取。）
- 同时，您也可以访问 `show proc` 语句中的 `CompactionStatus` 列中的 URL（在浏览器打开即可）来查看更具体的版本信息，来检查具体丢失的是哪些版本。  
如果长时间没有自动修复，则需要通过 `show proc "/cluster_balance"` 语句，查看当前系统正在执行的 `tablet` 修复和调度任务。可能是因为有大量的 `tablet` 在等待被调度，导致修复时间较长。可以关注 `pending_tablets` 和 `running_tablets` 中的记录。
- 更进一步的，可以通过 `admin repair` 语句来指定优先修复某个表或分区，具体可以参阅 `help admin repair`。

如果依然无法修复，那么多副本的情况下，我们使用 `admin set replica status` 命令强制将有问题的副本下线。具体可参阅 `help admin set replica status` 中将副本状态置为 `bad` 的示例。（置为 `bad` 后，副本将不会再被访问。并且会后续自动修复。但在操作前，应先确保其他副本是正常的）。

### Not connected to 192.168.100.1:8060 yet, server\_id=384

在导入或者查询时可能遇到这个错误。如果您去对应的 BE 日志中查看，也可能会找到类似错误。这是一个 RPC 错误，通常有两种可能：

- 对应的 BE 节点宕机。
- rpc 拥塞或其他错误。

如果是 BE 节点宕机，则需要查看具体的宕机原因。此处针对 `rpc` 拥塞的问题给出建议：

一种情况是 `OVERCROWDED`，即表示 `rpc` 源端有大量未发送的数据超过了阈值。BE 有两个参数与之相关：

- `brpc_socket_max_unwritten_bytes`：默认 1GB，如果未发送数据超过这个值，则会报错。可以适当修改这个值以避免 `OVERCROWDED` 错误。（但这个治标不治本，本质上还是有拥塞发生）。
- `tablet_writer_ignore_eovercrowded`：默认为 `false`。如果设为 `true`，则会忽略导入过程中出现的 `OVERCROWDED` 错误。这个参数主要为了避免导入失败，以提高导入的稳定性。

第二种是 `rpc` 的包大小超过 `max_body_size`。如果查询中带有超大 `String` 类型，或者 `bitmap` 类型时，可能出现这个问题。可以通过修改以下 BE 参数规避：`brpc_max_body_size`：默认 3GB。