腾讯云 | TencentDB 腾讯云数据库

PostgreSQL的云原生之路

基于分级存储与Serverless 结合的新一代数据库畅想

窦贤明
PostgreSQL/TDSQL-C PostgreSQL 产品研发负责人
腾讯云数据库专家工程师

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

01

云时代的数据库

数据库新篇章

02

云原生进化

存算分离

03

Serverless

极致弹性演进

04

分级存储

将存算分离进行到底

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

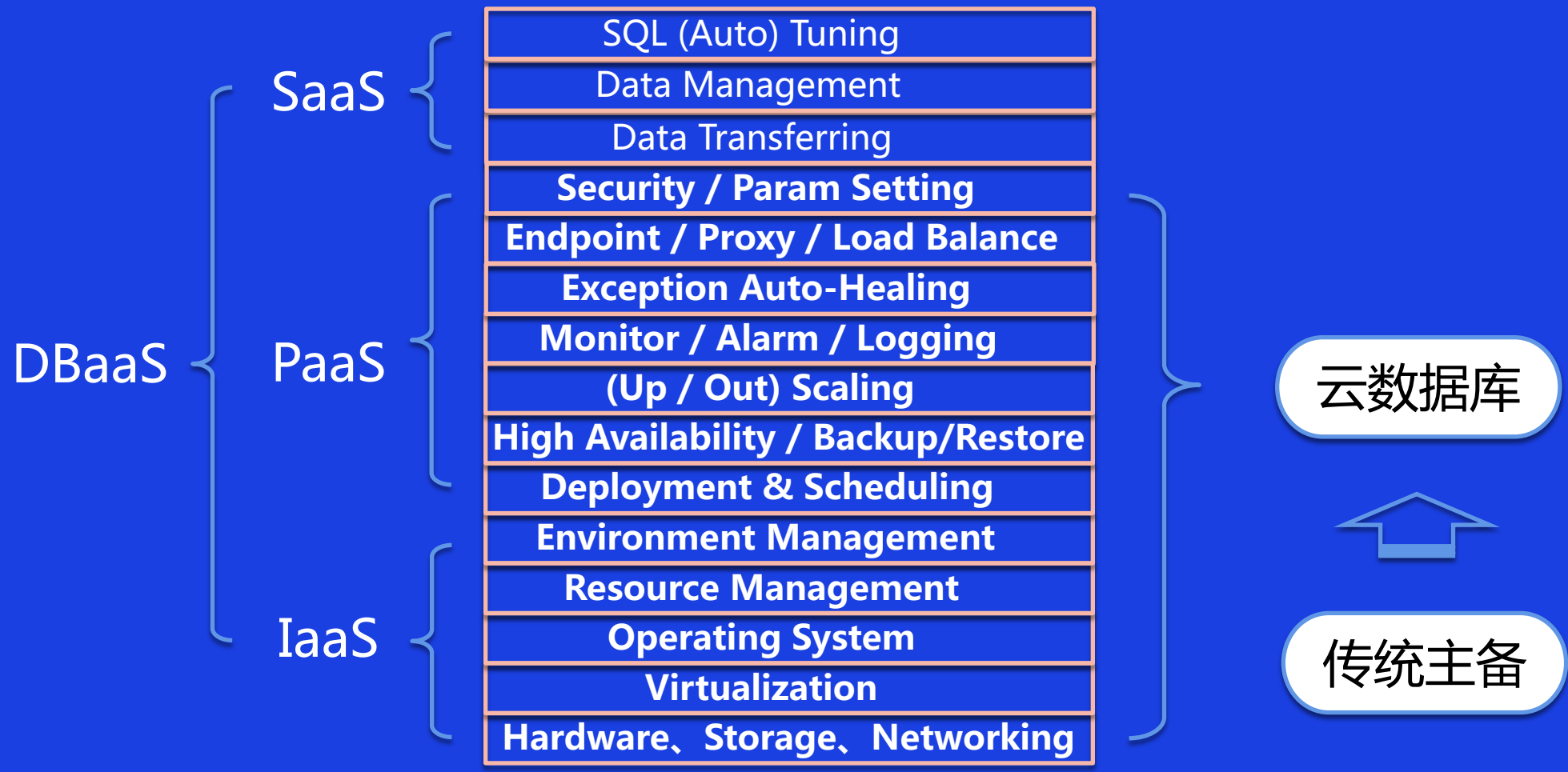
01

云时代的数据库

数据库新篇章

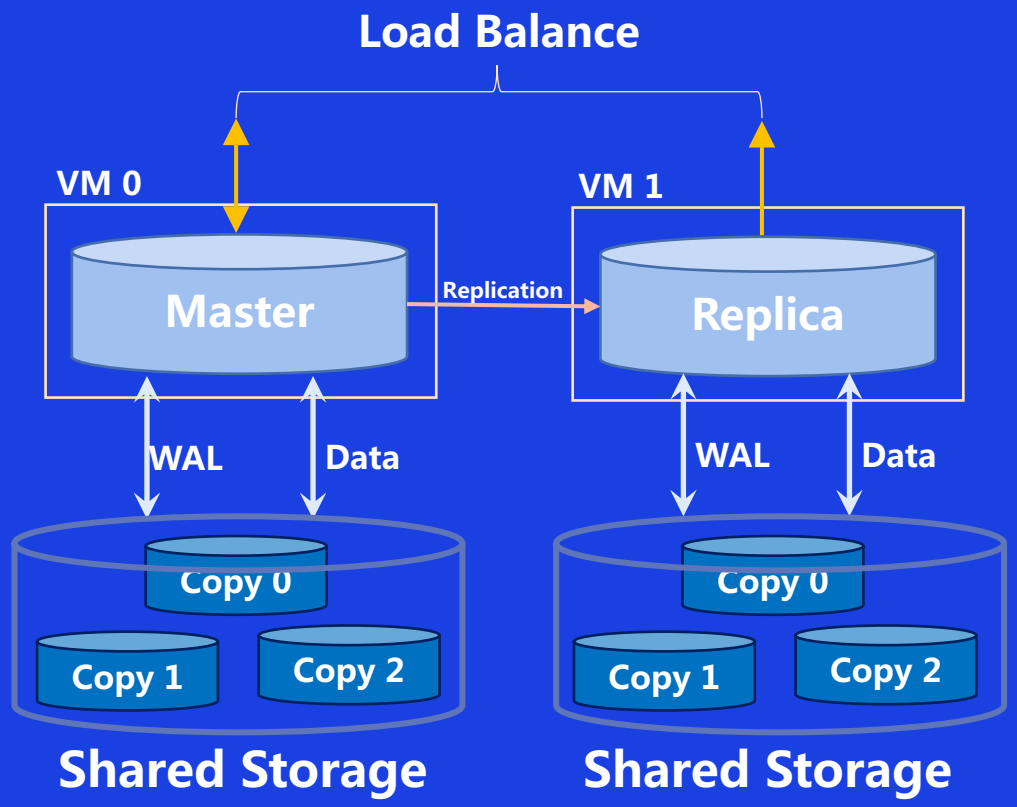
版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——TencentDB for PostgreSQL



版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——TencentDB for PostgreSQL



相比传统数据库：

- 成本更低
- 扩容能力更好
- 体验更好
- 成本核算的模式 从CAPEX转化为OPEX

主备方式：

- 不同节点的存储各自独立
- 主备节点间通过数据流复制保证数据一致
- 主库故障则切换到备库
- 可用性与可靠性的选择问题

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——TencentDB for PostgreSQL

云数据库未解决的问题：

- 存储成本，（3 + N * 3）份数据存储，N 为 RO 节点数量；WAL同理
- （高吞吐数据处理）网络成为瓶颈，共享存储侧有大量网络浪费
- RO建设成本高
- 弹性能力不够极致
- HA切换问题，可用性、可靠性的取舍

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

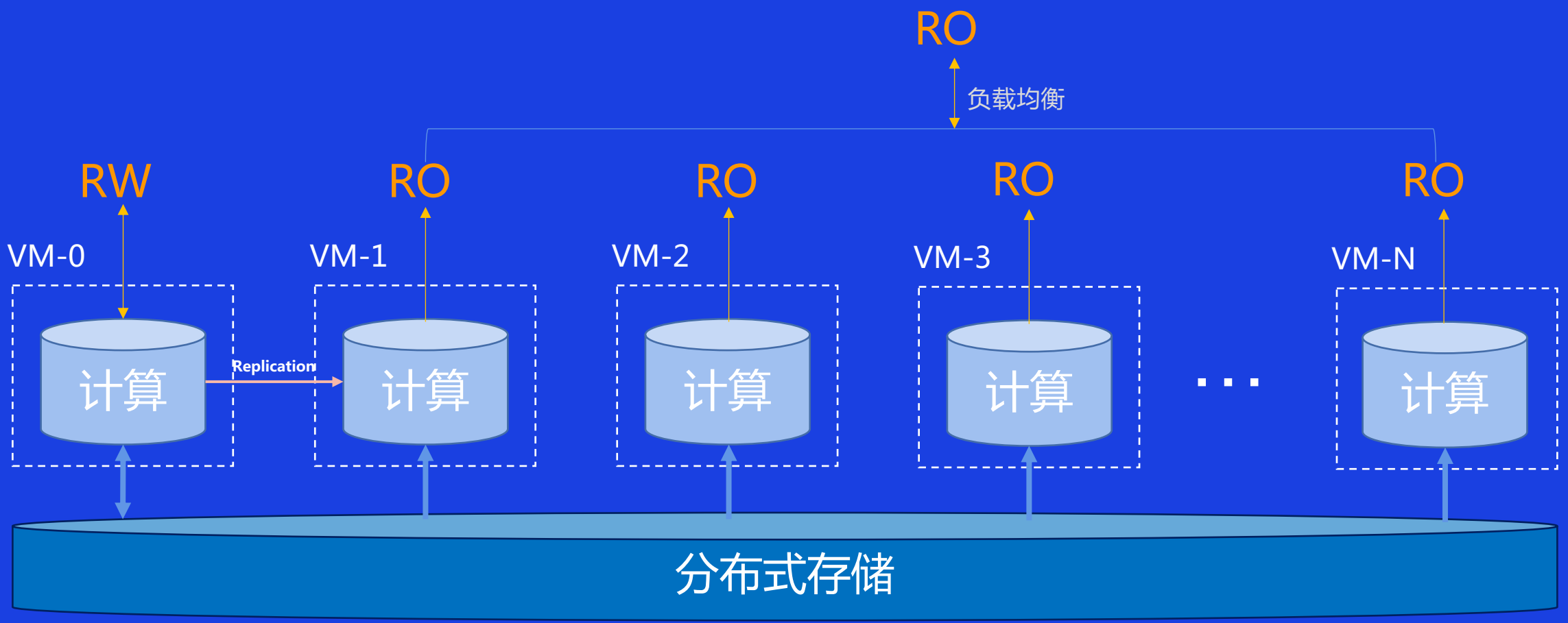
02

云原生进化

存算分离

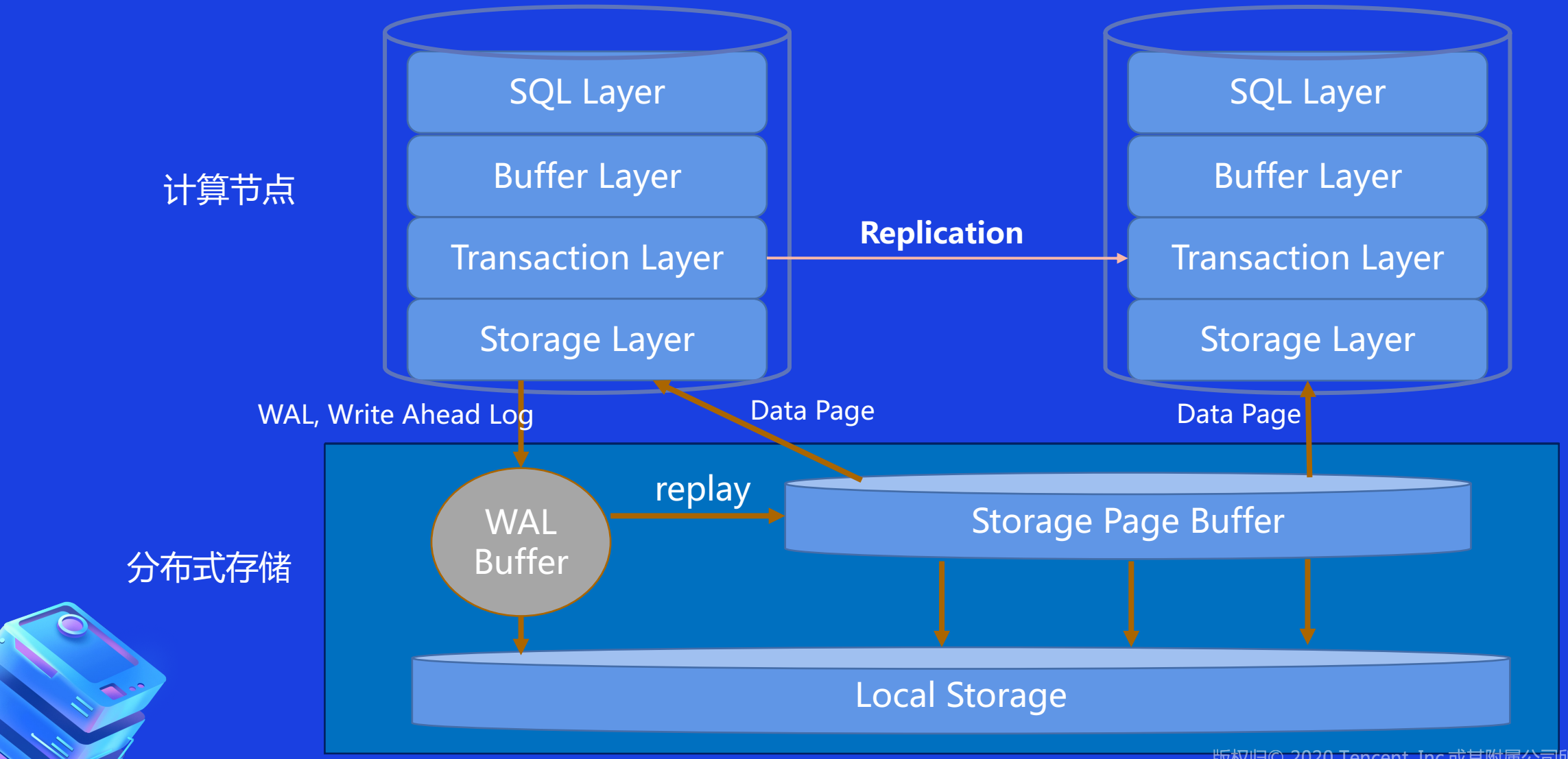
版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——TDSQL-C PostgreSQL



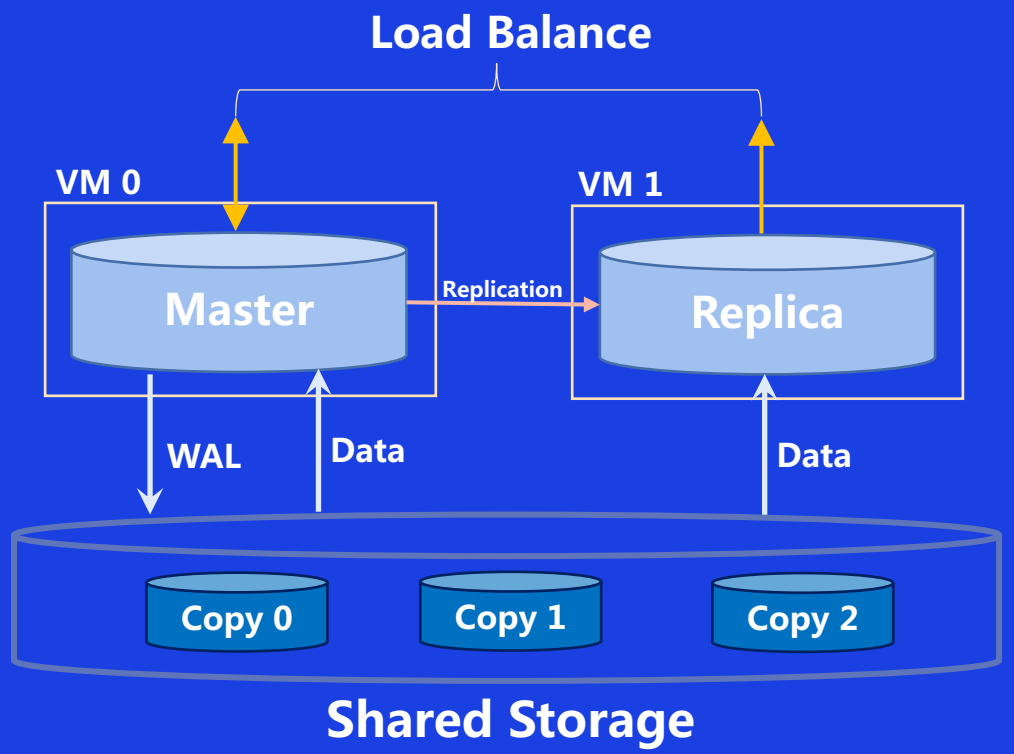
版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——TDSQL-C PostgreSQL



版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——TDSQL-C PostgreSQL



云原生数据库的优势：

- 存储、计算分离，各自动态扩缩容、分别计费，更优弹性
- 计算节点秒级拉起，保证可用性
- 状态持久化于分布式存储中，保证数据可靠性
- 总体存储成本更低，更灵活成本控制
- 存算分离，各自技术迭代，充分发挥硬件潜力

TDSQL-C PostgreSQL：

- Master 和 RO 基于一份数据，放在共享存储
- Master 仅将WAL写入共享存储、不写 **数据页**
- RO 从共享存储中读取所需 **数据页**，无须写存储
- RO 从主库接收 WAL、缓存中重放，保持缓存最新
- 共享存储接收并重放 WAL，实现存储节点上**数据页**的修改
- 存储层以 Page 为单位维护数据

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——TDSQL-C PostgreSQL

What is next ?

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

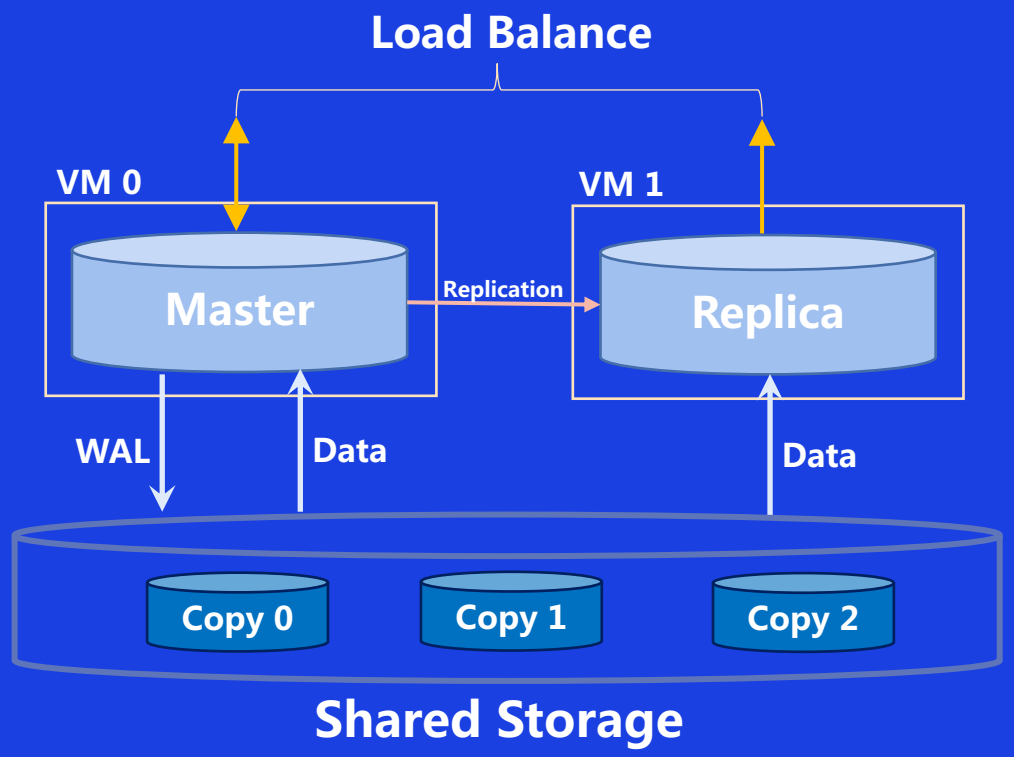
03

ServerlessDB

极致弹性

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——ServerlessDB PostgreSQL



云的本质：

- 弹性
- 自动化

极限演进思考：

- Master、Replica 秒级拉起，**极致弹性**
- 业务有高峰、低峰，甚至没有峰
- 计费，以实例为单位，时间周期为分钟级
- 数据库的用户为 应用和 DBA，只关心：
  - 地址
  - 计费
  - 运维

假设：

- 不再有实例，只是一个 Endpoint
- 不再有时间周期，仅计费实际使用的资源x时长
- 不再关心运维，全程自动化

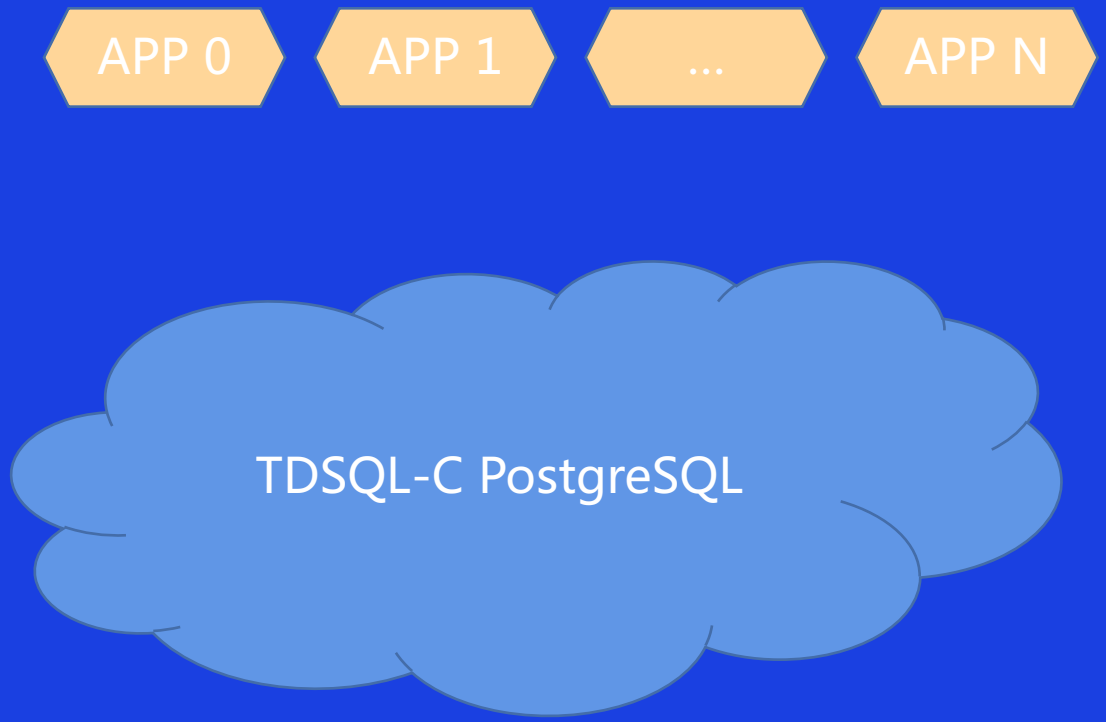
版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——ServerlessDB PostgreSQL

产品形态：

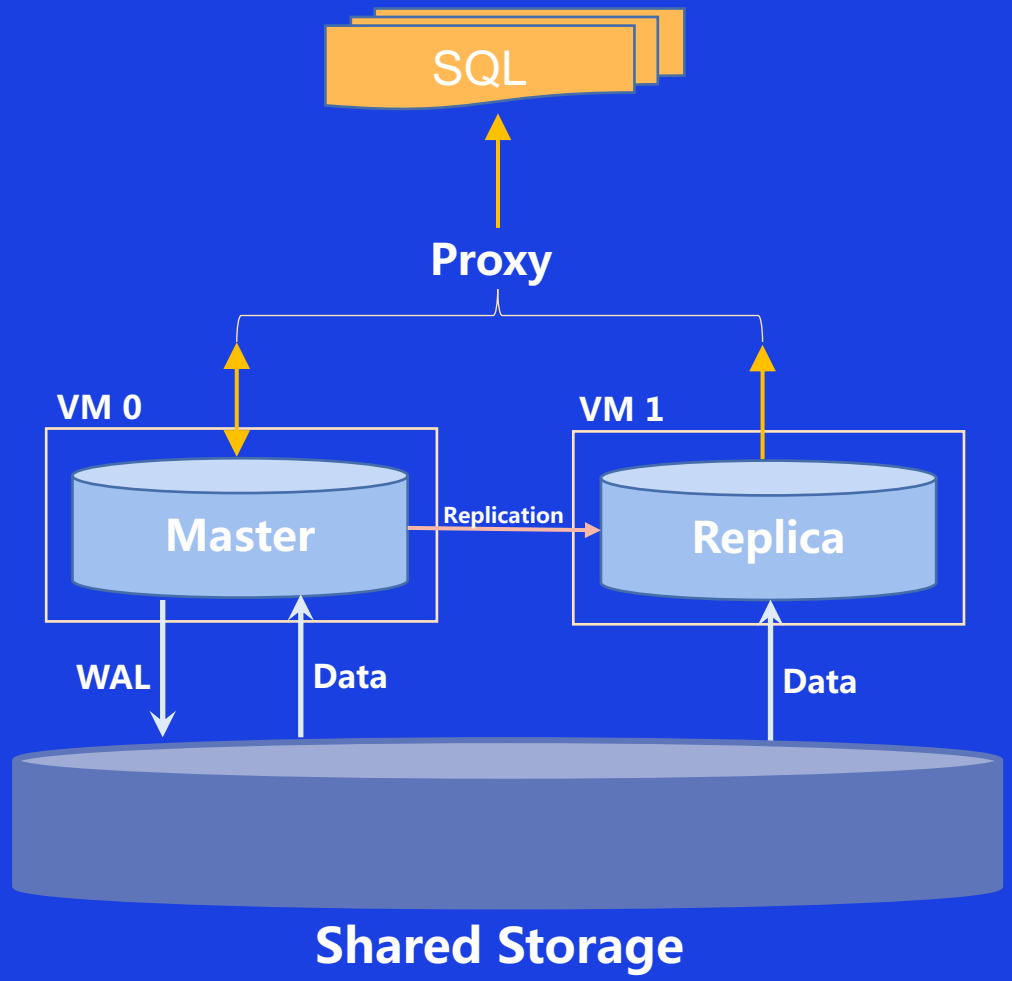
- 存储、计算分别计费，且粒度更细
- 计算节点，**不用不计费、用多长计多少**
- 计算节点，依据负载自动扩缩容
- 存储空间，用多少计多少
- 全程自动化、无需“人工”（或脚本）干预

- 用户只需关心：
  - 访问地址
  - 计费
  - 业务周期



版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——ServerlessDB PostgreSQL



初始状态：

- 存储存在、地址存在
- 计算节点不存在
- **没有业务**

SQL 运行时：

- 存储存在
- 计算节点被拉起
- **业务被执行**

SQL 运行结束后一段时间（数秒）：

- 存储存在
- 计算节点被关闭，不再计费

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

云时代的数据库——ServerlessDB PostgreSQL

Again, what is next ?

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

03

分级存储

将存算分离进行到底

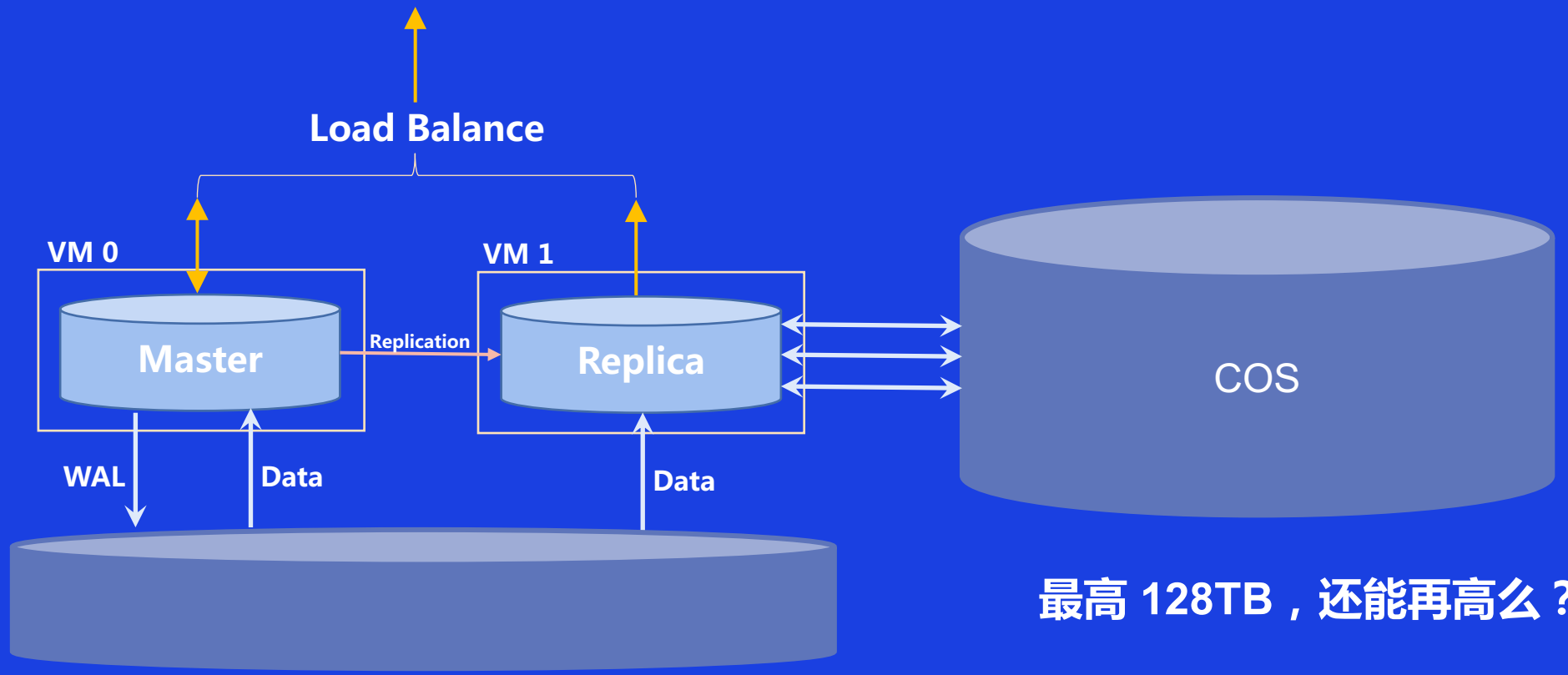
版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

分级存储

存算分离：计算做轻、存储做重

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

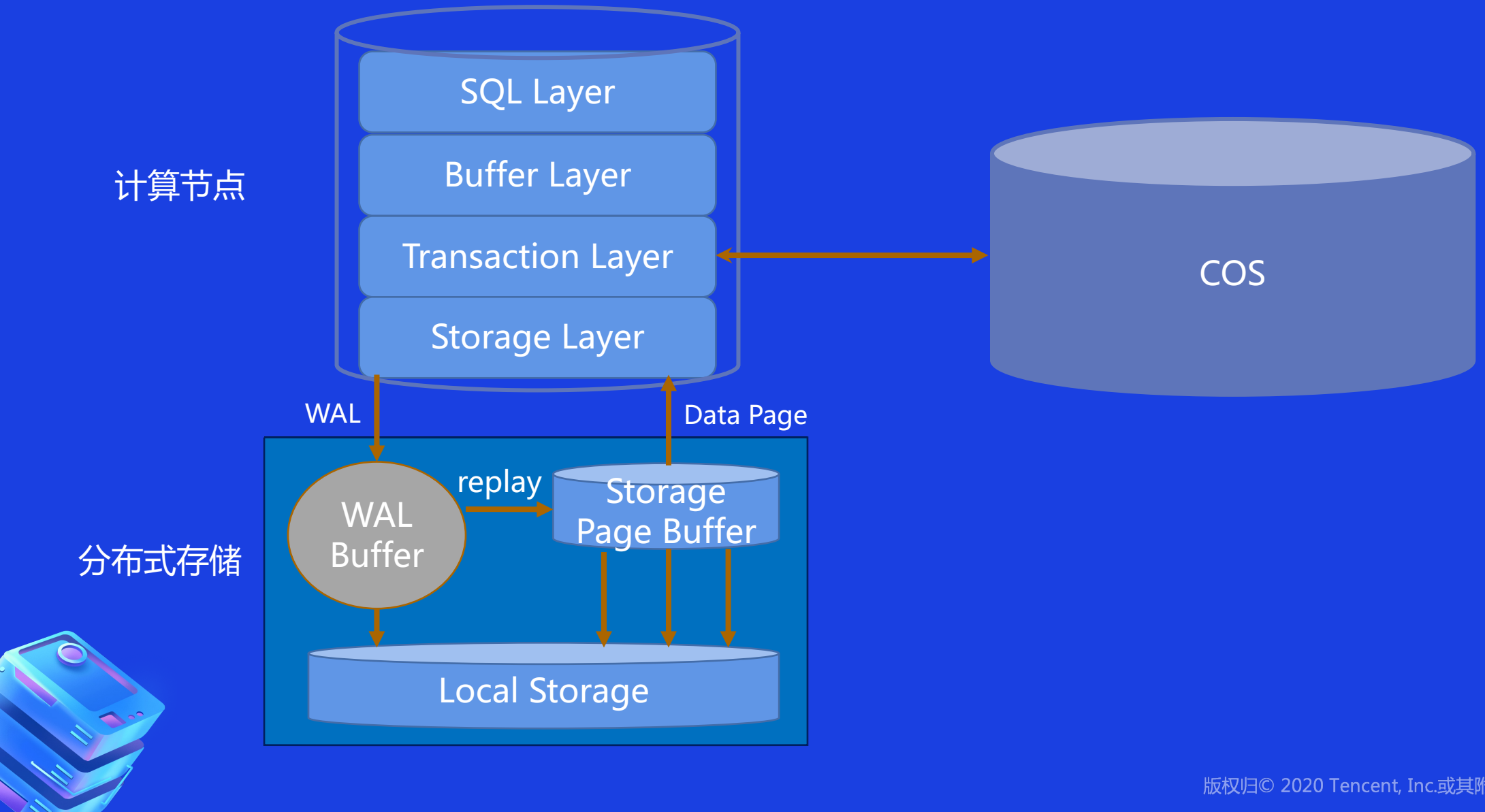
分级存储



最高 128TB，还能再高么？

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

分级存储



版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

分级存储

```
CREATE SERVER cos_server FOREIGN DATA WRAPPER cos_fdw OPTIONS(
                host 'xxxxxx.cos.ap-nanjing.myqcloud.com’,
                bucket 'xxxxxxxx’,
                 id 'xxxxxxxx’,
                key 'xxxxxxxxxx’
);
CREATE FOREIGN TABLE multi_csv (
                word1 text OPTIONS (force_not_null 'true’),
                word2 text OPTIONS (force_not_null 'off’)
) SERVER cos_server OPTIONS (
                filepath '/a.csv,/b.csv,/c.csv.2’,
                format 'csv’,
                null 'NULL’
);

postgres=# EXPLAIN SELECT * FROM multi_csv;
                              QUERY PLAN
-----------------------------------------------------------------------------
Foreign Scan on multi_csv (cost=0.00..1.20 rows=2 width=128)
                Foreign COS Url: https://xxxxxxxxxx.cos.ap-nanjing.myqcloud.com
                Foreign COS File Path: /a.csv,/b.csv,/c.csv.2
                Foreign each COS File Size(Bytes): 15,172,86
                Foreign total COS File Size(Bytes): 273
(5 rows)
```

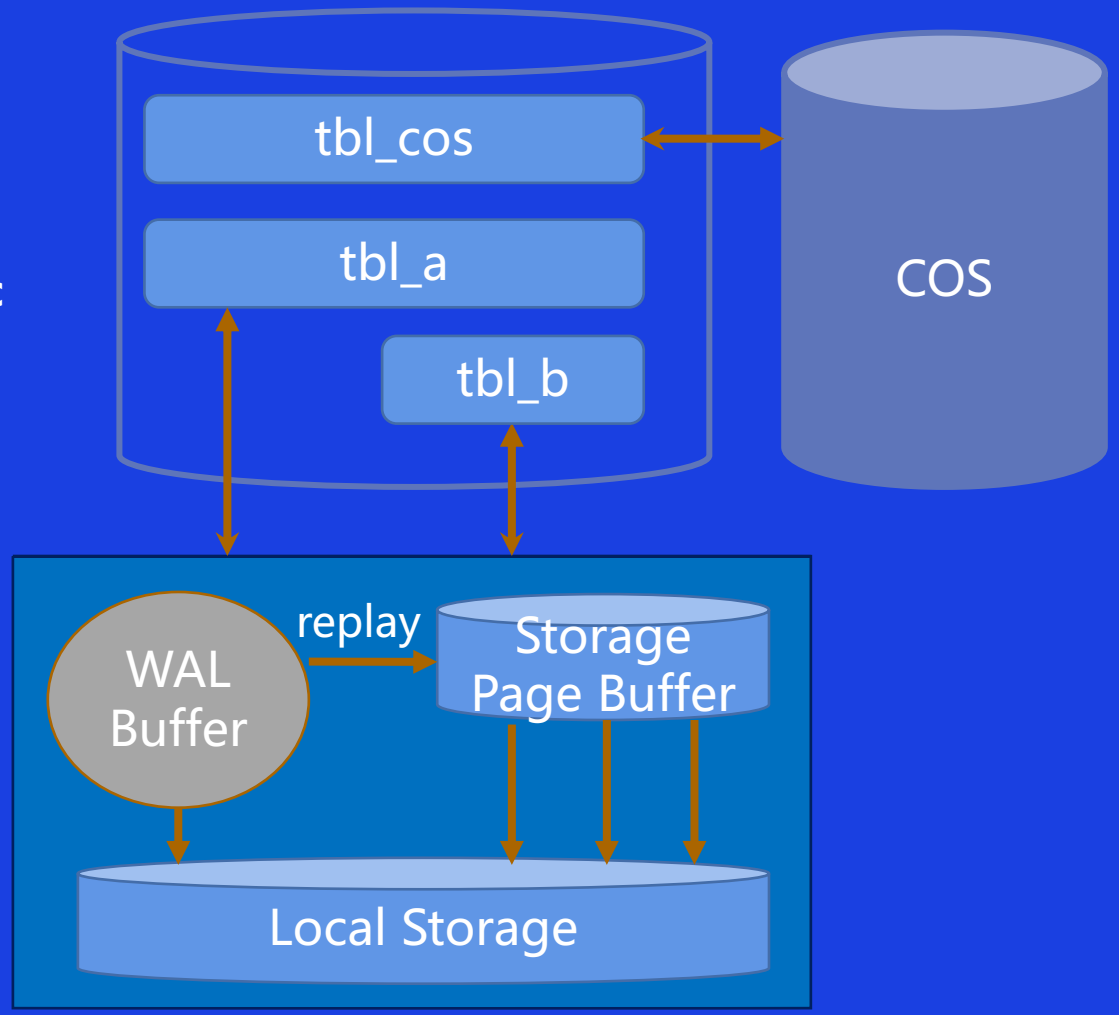
版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

分级存储

APP 0

```
SELECT
        a.id, b.name, c.value
FROM
        tbl_a a, tbl_b b, tbl_cos c
WHERE a.id = b.a_id and b.id = c.b_id
...
```

By Serverless, Or Cluster



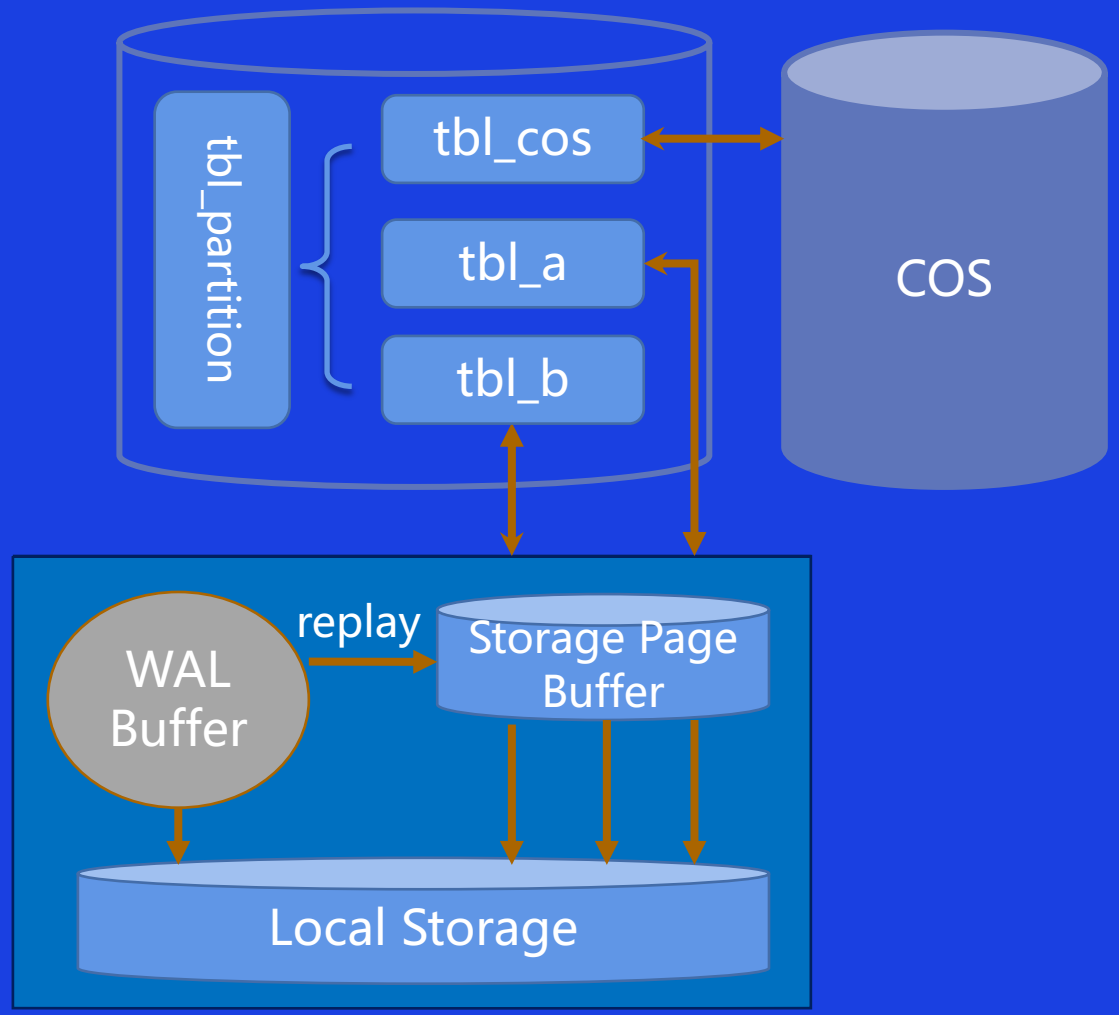
版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

分级存储

APP 1

```
SELECT *
FROM
        tbl_partition
WHERE
        date > ' 2022.04.30'
```

By Serverless, Or Cluster



版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利

版权归© 2020 Tencent, Inc.或其附属公司所有 保留所有权利