



第十一届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2020

架构革新 高效可控



北京国际会议中心 | 2020/12/21-12/23

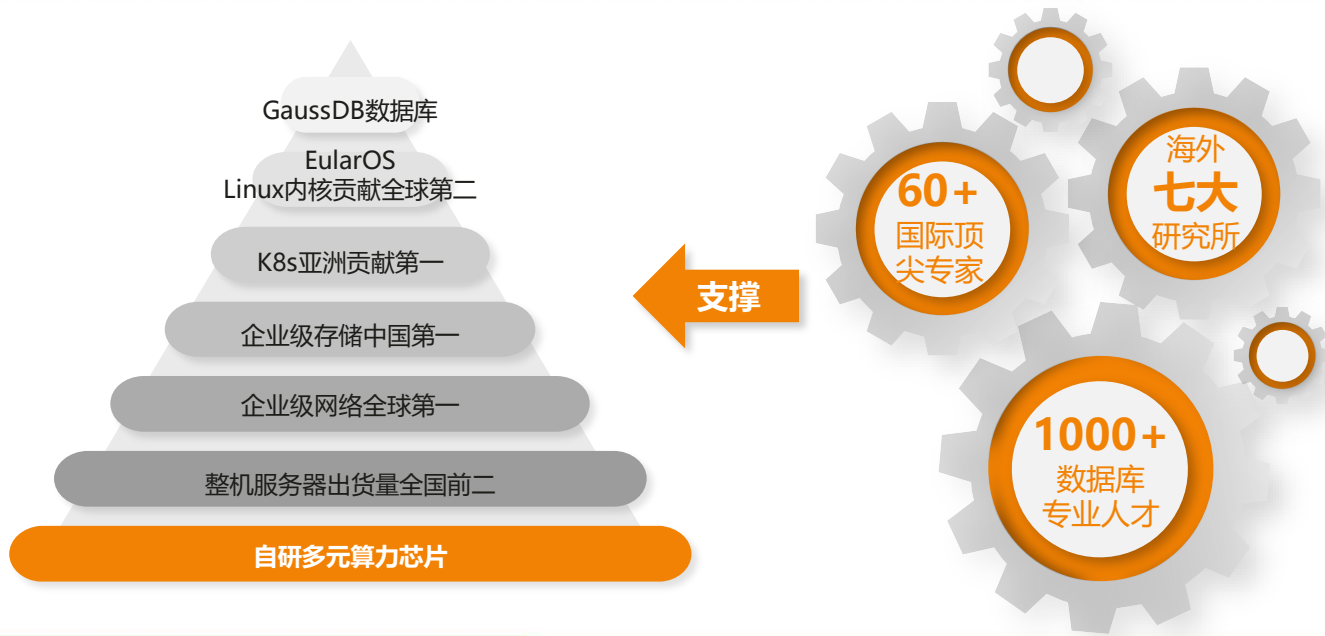
华为云数据库GaussDB 使能企业数字化升级的实践与探索

彭立勋 | 华为云数据库技术专家



立足华为云原生全栈能力

华为公司在多年各行业的企业级客户优秀实践中，锤炼总结了贯穿**多元化芯片、整机服务器、高速存储、新一代网络及企业级软件**等领域的整合优化经验，并将其持续投入在数据库领域中。



华为全球布局1000+数据库专业人才，持续投入超10年



8大全球研究所从事基础研究，**10+**年数据库领域技术积累，**1000+**数据库专项人才，**40000+**数据库集群



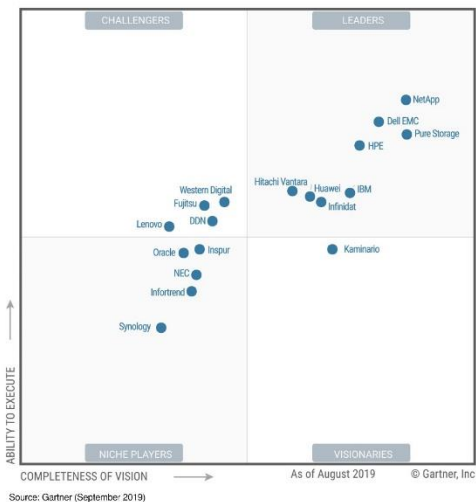
SIGMOD 2020: 数据库顶级会议

The screenshot shows the ACM Digital Library interface. The top navigation bar includes 'Magazines', 'Proceedings', 'Books', 'SIGs', 'Conferences', and 'People'. A secondary bar lists 'Conference', 'Proceedings', 'Upcoming Events', 'Authors', 'Affiliations', and 'Award Winners'. The breadcrumb trail reads: 'ences > MOD > Proceedings > SIGMOD '20 > Taurus Database: How to be Fast, Available, and Frugal in the Cloud'. The article title is 'Taurus Database: How to be Fast, Available, and Frugal in the Cloud'. The authors listed are Alex Depoutovitch, Chong Chen, Jin Chen, Paul Larson, Shu Lin, Jack Ng, Wenlin Cui, Qiang Liu, Wei Huang, Yong Xiao, and Yongjun He, with a link to 'Authors Info & Affiliations'. The publication information states: 'Publication: SIGMOD '20: Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data • June 2020 • Pages 1463–1478 • <https://doi.org/10.1145/3318464.3386129>'. At the bottom of the screenshot, the full DOI URL is repeated in orange text: <https://dl.acm.org/doi/abs/10.1145/3318464.3386129>.



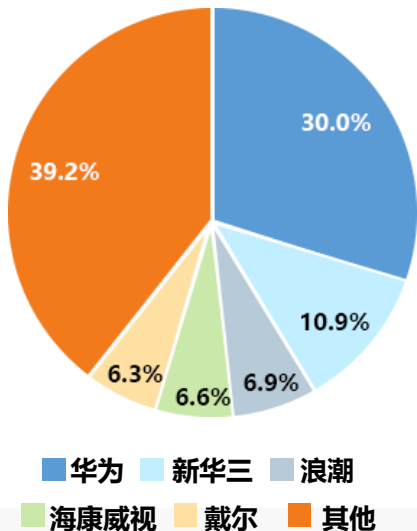
基于华为企业级分布式存储底座

Figure 1. Magic Quadrant for Primary Storage



2020年上半年中国企业级存储市场

前五大厂商市场份额

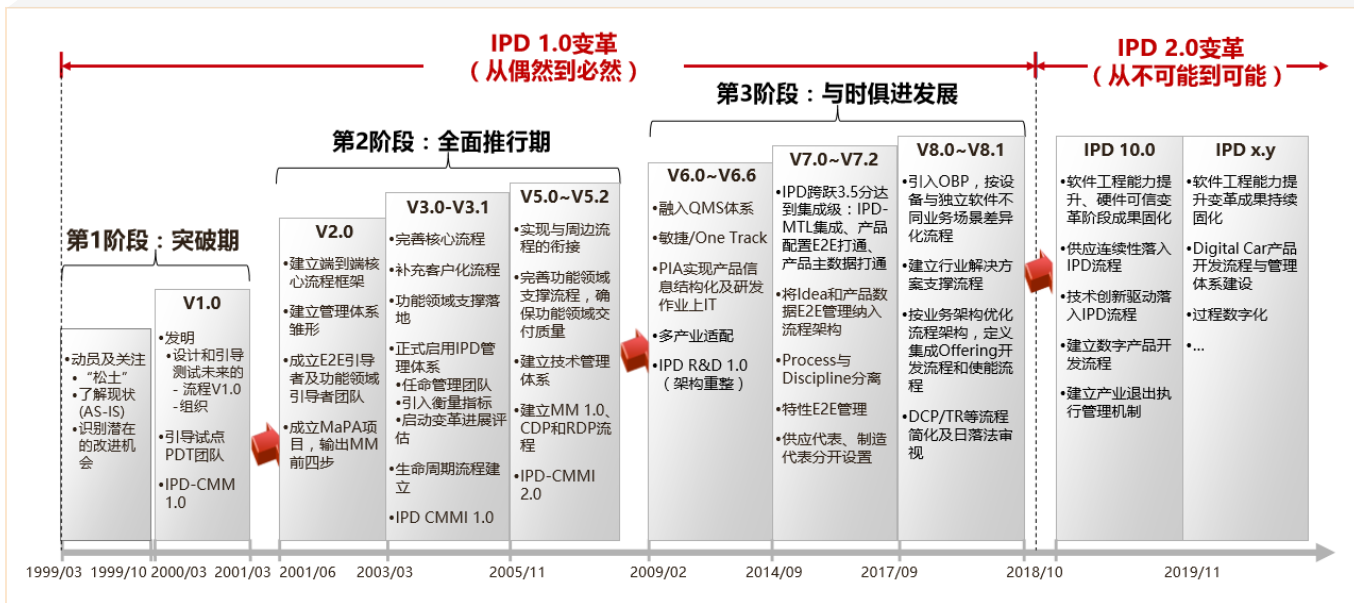


数据库的核心是**数据安全**，数据安全的保障是**存储可靠**。

华为云GaussDB数据库存储基于华为自研存储管理软件DFV，华为存储连续4年Gartner象限领导者地位，**20年存储领域积累**，中国区市场长期第一，累计管理着数百万在线硬盘，市场证明华为存储底座的全场景可靠性。

依托电信级质量体系

华为从1999年起引入IPD并结合自身发展与业界实践持续优化



华为云数据库秉承华为电信级交付质量体系，结合云原生场景，打造业界顶尖的软件质量。覆盖硬件/软件/业务类场景，包含超过50万测试用例，代码覆盖率达90+%。

华为云数据库的新里程碑

Magic Quadrant

Figure 1: Magic Quadrant for Cloud Database Management Systems

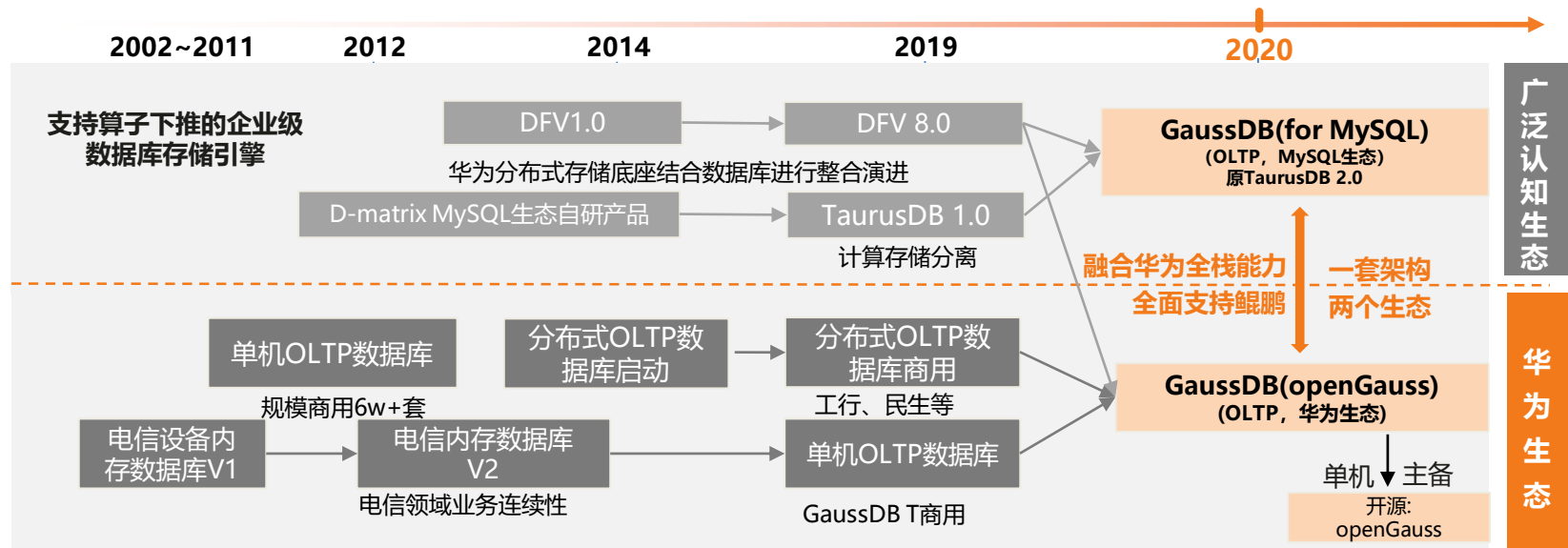


Source: Gartner (November 2020)

日前，国际知名调研机构Gartner发布2020年全球云数据库魔力象限报告（《Magic Quadrant for Cloud Database Management Systems》），华为云凭借 GaussDB 系列数据库产品入选魔力象限。此次入选，意味着华为云在云数据库领域赢得了市场的广泛认可。



长期投入数据库：一套架构，两个生态，坚持开放



长期投入数据库的战略不变：

- 全面布局数据库产业：聚焦自研内核、华为生态及拥抱开放生态；
- 先进的数据库架构：新一代分布式理念根植产品，扩展传统RDBMS范畴，底层集成资源调度管理，为云化演进路线打下基础。

云数据库以市场为导向，生态开放，架构先进， 打造全场景云服务

自研

开源

关系型数据库服务

GaussDB(for MySQL)

RDS for
PostgreSQL

RDS for
MySQL

GaussDB (openGauss)

RDS for
SQL Server

非关系型数据库服务

DDS
社区版

- GaussDB (for Mongo)
- GaussDB (for Influx)
- GaussDB (for Cassandra)
- GaussDB (for Redis)

数据库工具服务

UGO

数据库与应用迁移服务

DRS

数据迁移复制服务

DAS

数据管理服务

DDM

分布式数据库中间件

GaussDB数据库引擎的3层架构：生态兼容，层次解耦，数据融合

SQL Interface Layer: 生态

- 每个数据Server引擎的实现，代表独特的生态
- 没有一个数据库可以做到100%兼容其它数据库
- NoSQL领域甚至某种类型数据，一个引擎，如：MongoDB，Cassandra等
- 基于一个生态，而不是重写一个生态+兼容，是最佳的选择

Index Layer: 数据组织和访问 不同引擎需要不同的数据组织和存储语义

- 不同引擎的数据库数据格式不同，存储模式不同，需要不同的插件
- 不同的引擎无法统一；

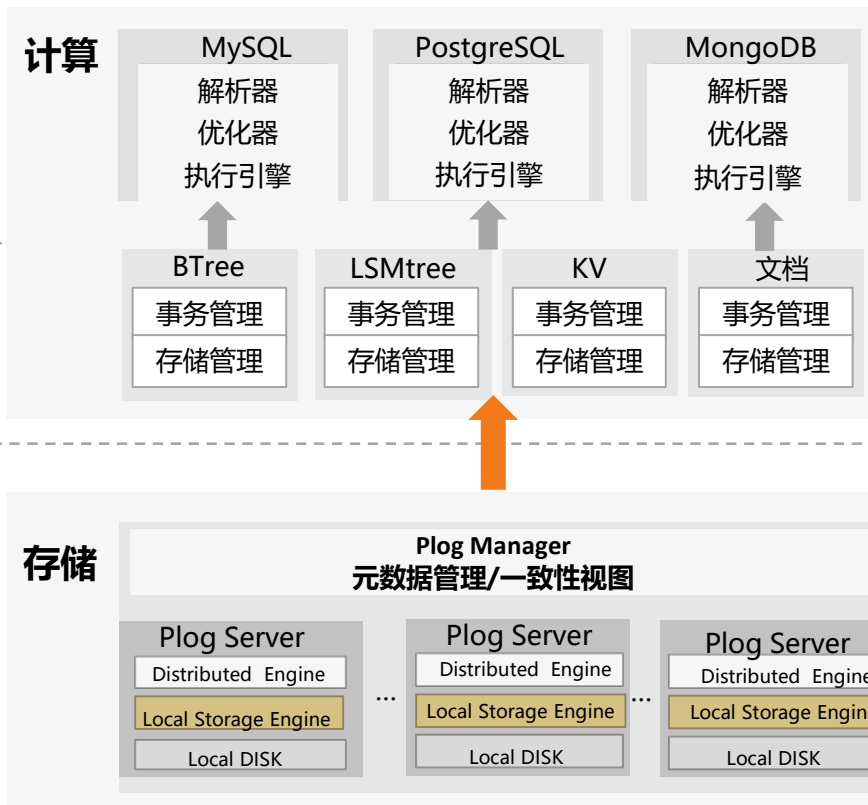
存储计算分离

Storage layer: 跨AZ，脱离语义的数据能力 提供基础分布式一致性可扩展存储能力

- 高效访问：融合网络、存储各种新技术。
- 分布式一致性：分布式管理，提供可扩展存储能力。
- 跨AZ
- 融合资源池

算子下推：

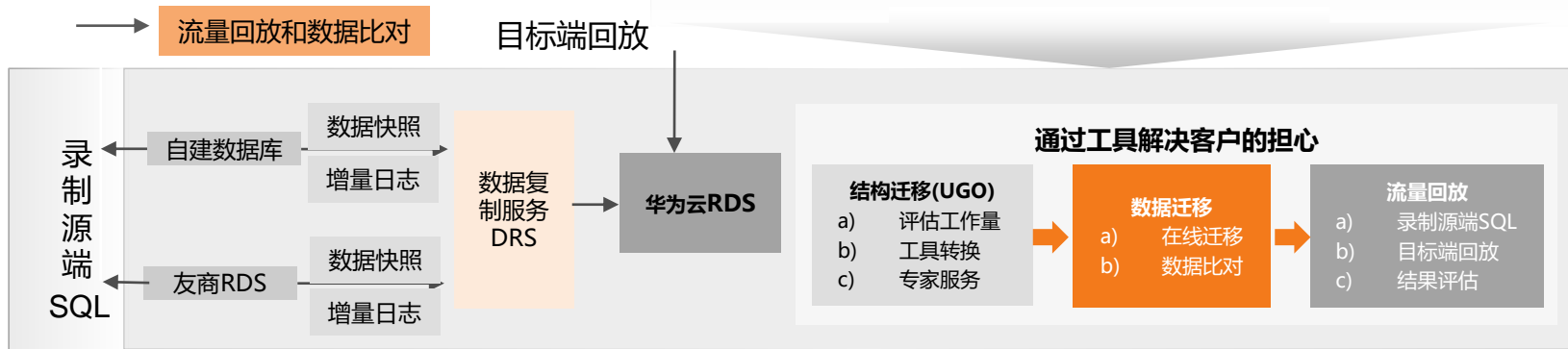
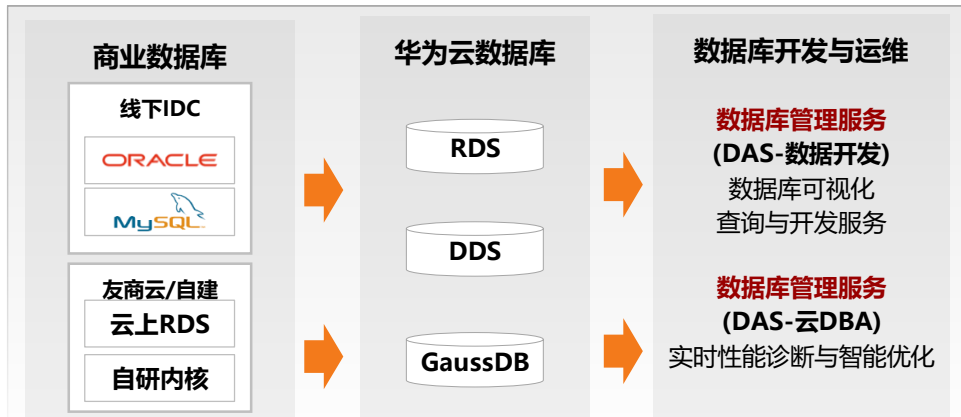
原生DFV上，支持Index Layer逻辑部分下推到存储，减少网络IO



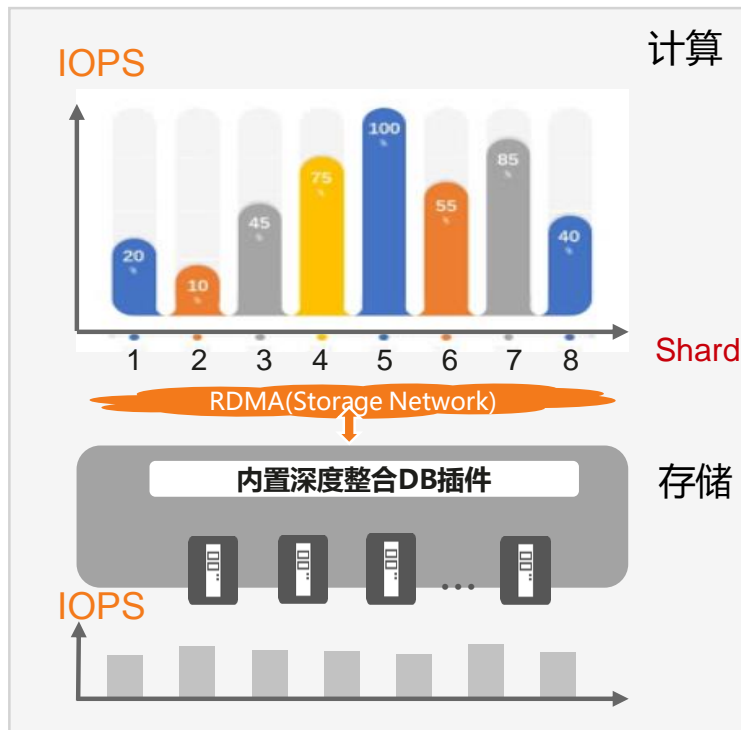
华为云数据库打造一站式的上云解决方案

客户，对交易型数据和系统，普遍的4个担心：

- ① 迁移改造工作量会不会很大？
- ② 迁移过程数据会不会丢？
- ③ 迁移过来是否兼容？
- ④ 迁移后表现是否与POC一致

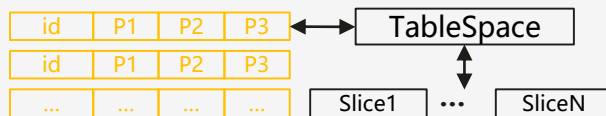


云原生存算分离架构解决热点数据问题



计算侧均匀打散页面

SliceMapping



MySQL的页面映射到slice (10G) , 打散IO

分布式存储侧均匀IO

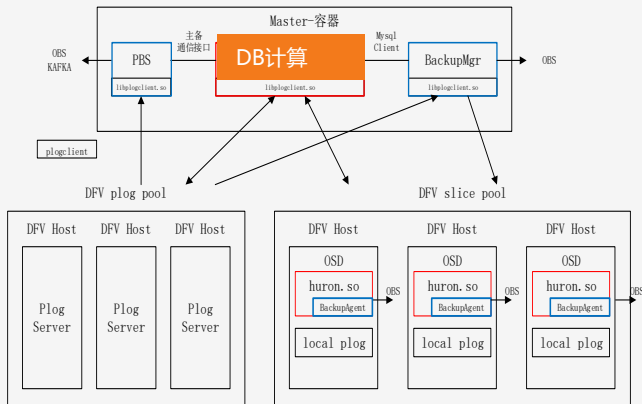
Plog Manager
元数据管理/一致性视图

Plog Server ... Plog Server Plog Server

分布式存储基于大小、IO分配算法均匀分布

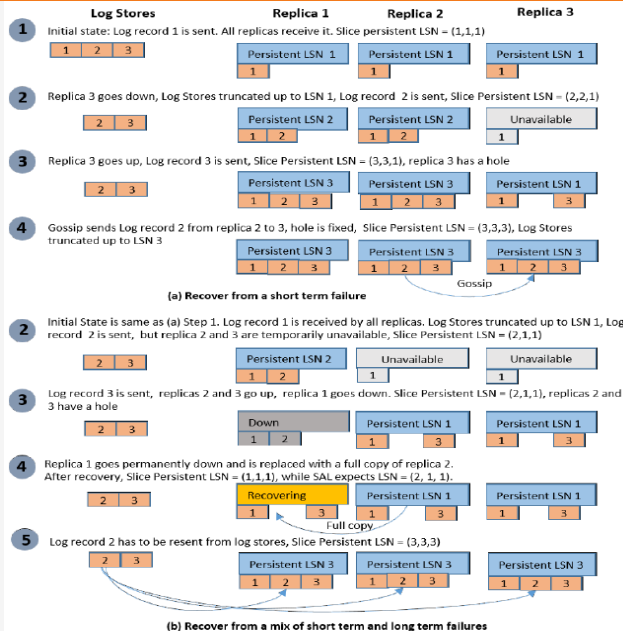
云原生存算分离架构数据0丢失，故障闪恢复

计算无状态



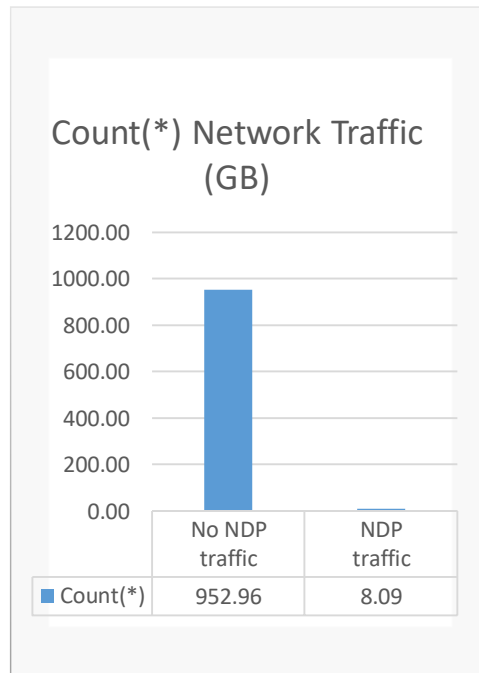
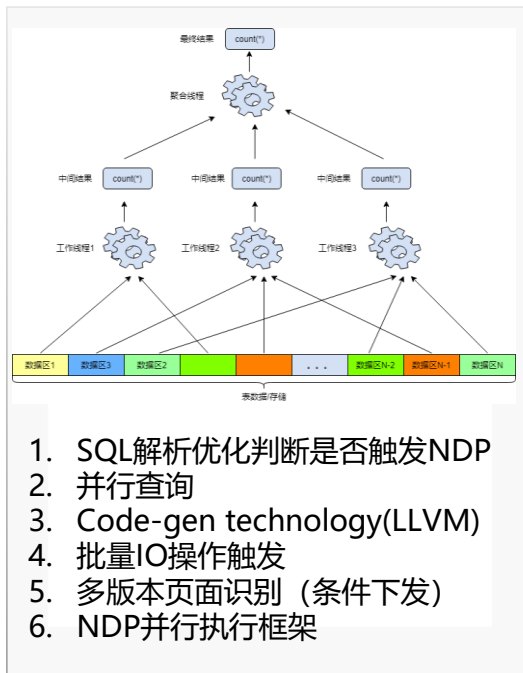
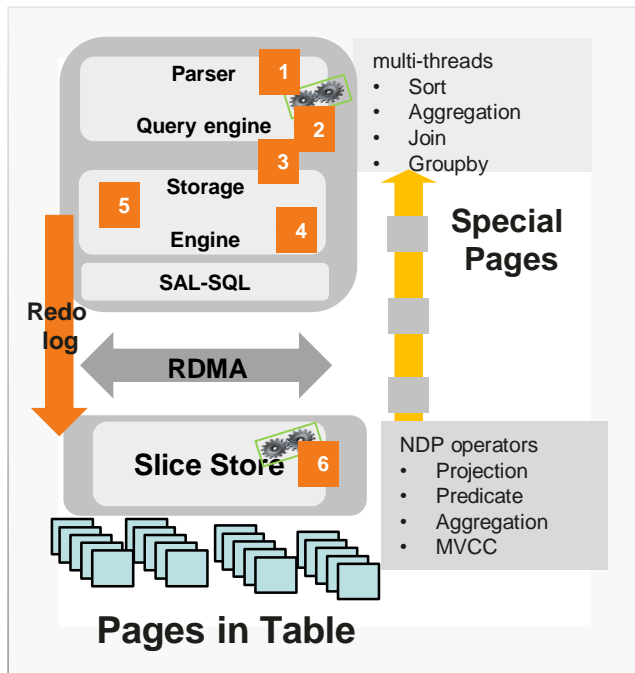
- 计算层无状态，RTO 10S
- 依托华为分布式存储 RPO = 0
- 数据库存储深度融合，强一致。
- 基于快照备份，分钟级备份性能

存储强一致



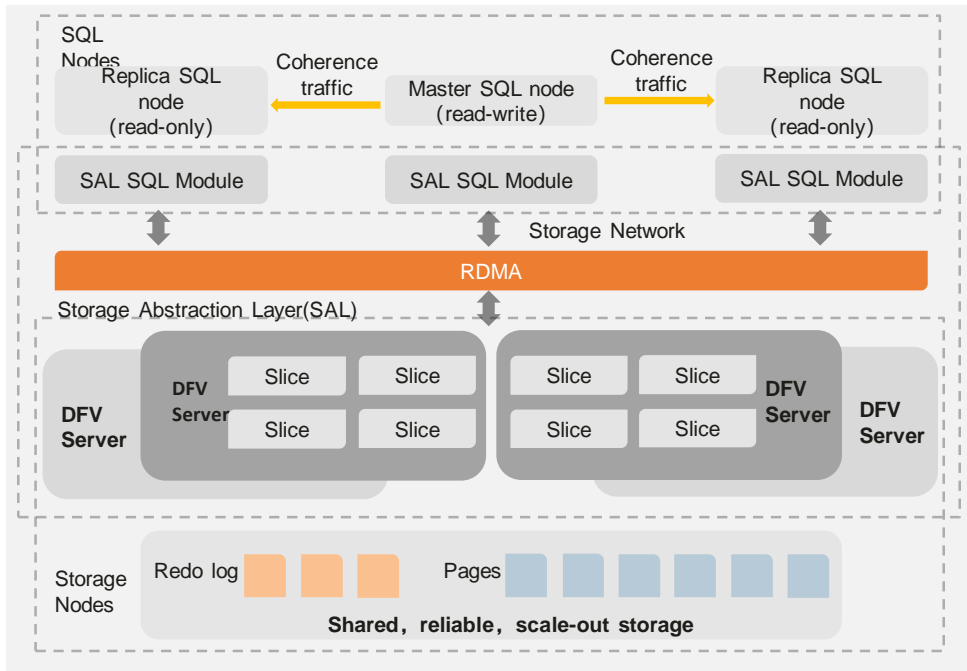
性能: PQ (并行查询) + NDP (近存储计算) 框架支持算子下推

GaussDB (for MySQL) 分布式版NDP architecture



GaussDB for MySQL: 一写多读, +DDM实现多写多读

兼容MySQL生态, 基于鲲鹏架构, 计算存储分离、云化架构的企业级云数据库



GaussDB for MySQL数据库核心优势

超高性能

- 传统MySQL的五倍以上
- 100W读
- 30W写

高扩展性

- 15只读副本, 128T存储

高可靠性

- 跨AZ部署, 数据三副本, 99.99%可用性

高兼容性

- MySQL兼容, 结合 DRS 从 MySQL 在线迁移

应用场景

金融、娱乐

- ✓ 高吞吐量
- ✓ 大数据量处理

手游、电商

- ✓ 高可用
- ✓ 弹性伸缩能力

电信、互联网

- ✓ 高可靠
- ✓ 异地容灾

案例：华为消费者云全面切换到GaussDB for MySQL

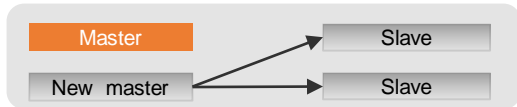
整体数据量PB级，单业务最大容量100T，900套数据库遍布全球

数据库架构演进之路：从HA4副本-> GaussDB for MySQL存算分离

准强一致



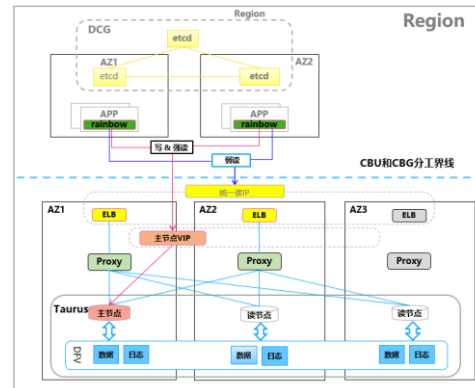
单点故障



AZ故障



- 物理存储从6->3
- 可靠性准强一致->强一致
- 性能：分布式存储动态均衡，避免单点热

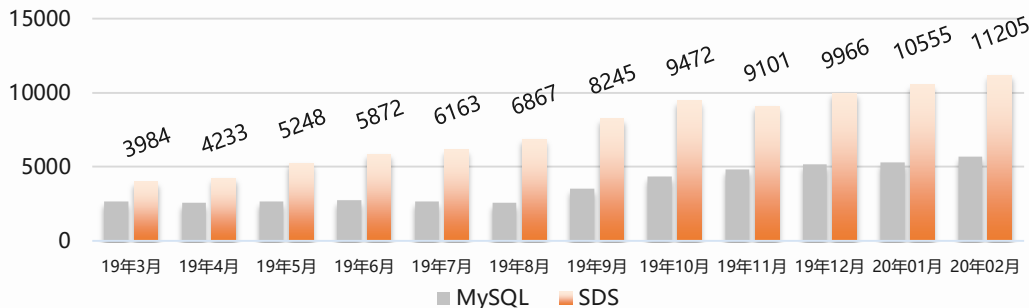


故障场景	概率	传统HA版可靠性	3副本强一致存算分离可靠性
磁盘	0.3% ~1	Raid保障	3副本保障，不仅有故障还有慢盘保障
物理机	5%	故障节点恢复小时级	故障节点恢复分钟级
网络	跨AZ高发抖动	降到异步或性能抖动	消费者业务7*24性能无1s抖动
机柜故障/机房/Region	天灾人祸,核心是电力	降级到主备，运行有丢失数据风险	折叠，2AZ，3副本，单点故障不降级。

消费者云主要业务场景

数据库规模增长非常迅猛，扩展性要求高

1. MySQL实例从2670->5698增长2.1倍
2. Cassandra节点从3984->11205增长2.8倍



业务模块	使用的数据库	数据场景	数据库规模
华为会员	MySQL	华为ID	58TB
视频	MySQL, Cassandra	评论、弹幕	43TB
音乐	MySQL, Cassandra	歌单、歌曲	40TB
运动健康	Cassandra	数据流水	550TB
云空间, 相册	MySQL, Cassandra	相册索引	690TB
应用市场	MySQL, Cassandra	账单	35TB

1. 消费者云开源MySQL后续全部转为GaussDB (for MySQL)
2. 消费者云Cassandra后续会有一部分转化为GaussDB (for Cassandra)
3. 消费者云Cassandra后续会有一部分转化为GaussDB (openGauss)
4. 消费者云线下有使用分库分表的业务，主要是做分库分表和连接收敛，后续会逐步迁移至为GaussDB (openGauss)

华为消费者云MySQL替代案例

开源MySQL全部迁移到GaussDB (for MySQL)

主要应用：HUAWEI会员、用户账单等

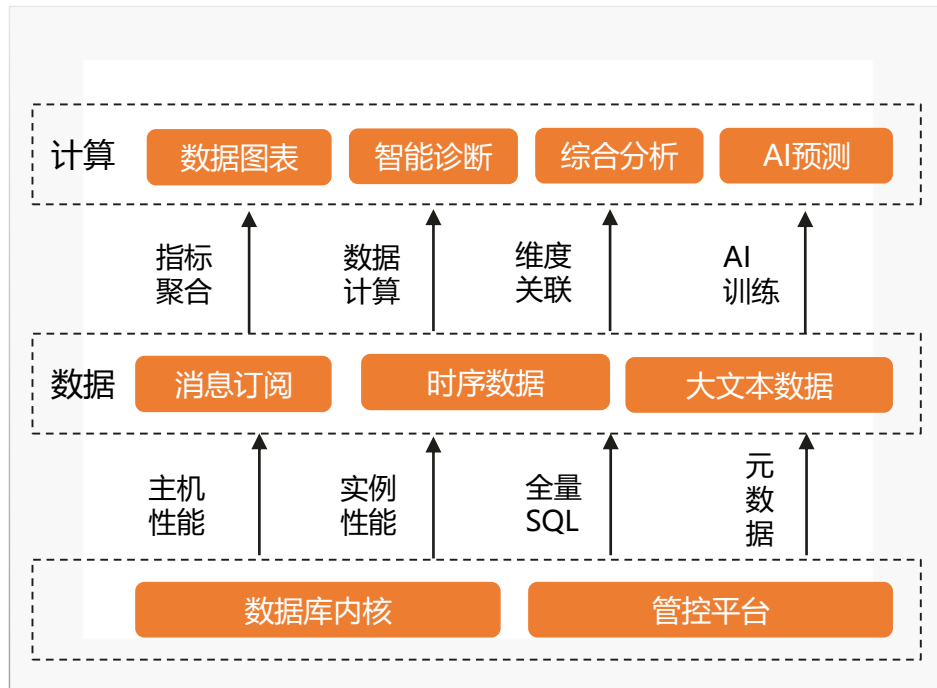
	2020年	2021年	2022年
用户订购关系表	套餐订购用户：10W	餐订购用户：30W	餐订购用户：60W
用户免费资源表	用户规模：10W 数据量：约1.8亿	用户规模：30W 数据量：约8.1亿	用户规模：60W 数据量：约21.6亿
账单表	同用户免费资源表，按需套餐计费周期为小时，账单是小时级出账	同用户免费资源表，按需套餐计费周期为小时，账单是小时级出账	同用户免费资源表，按需套餐计费周期为小时，账单是小时级出账
数据规模	150W订购用户，每个用户创建5个产品，每个产品购买30个服务，每个服务3个计费项，开通了按需计费（小时级账单），全年数据量约为：17.4W亿		

1、单库容量和性能都无法满足需求，需要分库。按照产品ID做Hash分片，由于用户量是逐步增长的有一个过程，初始上线时分库节点数为3，单库每月约为1.8亿数据

2、单库单表每年的数据量为21亿左右，需要按照日期做分表，采用按月分表的方式，每月单表数据量约为1.8亿（运营有按月查询的场景）

3、由于数据量过大，数据库中只保存最近3个月的热点数据，超过3个月的数据同步到BI HDFS存储，由BI提供3个月之前的运营报表，并将原始数据存档保存2年。

DAS:专业的数据库运维平台



实时会话&性能

历史性能趋势

慢SQL分析与诊断

全量SQL分析与诊断

实时分析

异常诊断

每日巡检

紧急Kill会话

实例参数智能推荐

DRS(Data Replication Service)华为云数据复制服务

企业级端到端**数据实时**流转平台级产品，集在线迁移、多活灾备、实时同步能力为一体。

全球2000+企业规模商用，遍布金融、政府、电信、互联网等行业

全球10000+任务，同城ms、异地s级同步速度



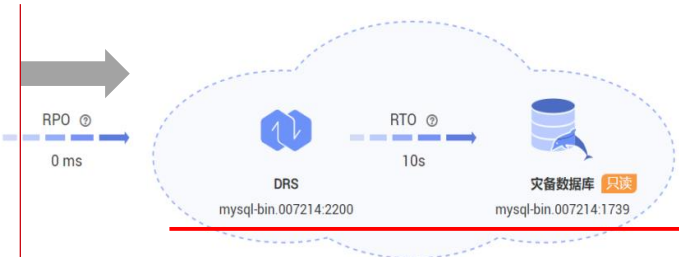
2019DTCC大奖



DRS支持不同云和不同系统之间的数据库的双活

基本信息 灾备进度 灾备监控 灾备对比 灾备日志 标签

本云数据库升主

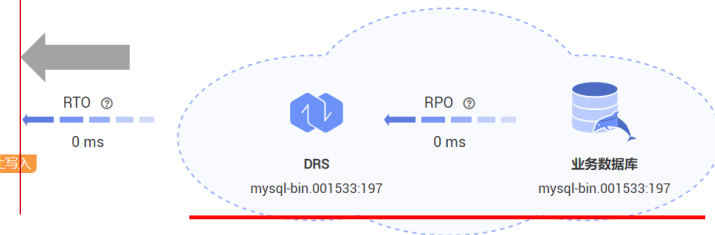


华为云侧

一键切换

基本信息 灾备进度 灾备监控 灾备对比 灾备日志 标签

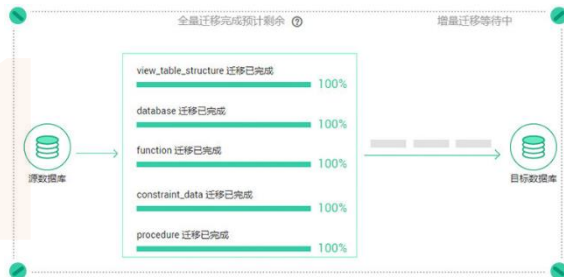
本云数据库降备



华为云为主

迁移和数据同步过程有详细的信息监控

一看宏观展示，实时了解迁移进展



宏观查看全量迁移对象的百分比进度，如迁移时长较多的表数据、表结构、表索引等。



二看表格信息，一目了然各项迁移指标进度

进度明细

迁移对象	对象数目	状态	已迁移数目	详细信息
procedure	0	迁移中	0	查看详情
routine	0	迁移中	0	查看详情
table_indexes	0	迁移中	0	查看详情
table	11	迁移中	0	查看详情
database	2	已完成	2	查看详情
view	0	迁移中	0	查看详情

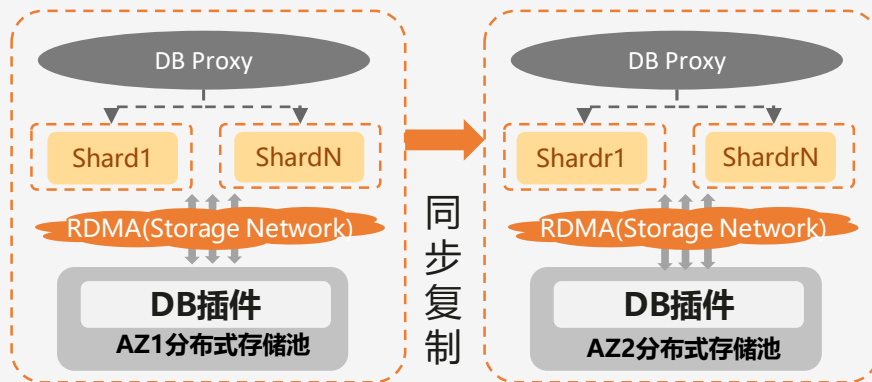
通过表格查看具体迁移对象的迁移进展，当“对象数目”和“已迁移数目”相等时，表示该对象迁移完成。

三看具体迁移对象进度

通过“查看详情”精细查看每个对象的迁移进度，当进度显示为100%时，表示迁移完成。

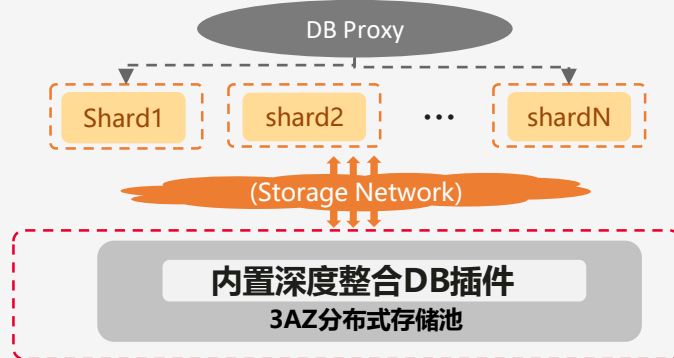
GaussDB for MySQL 分布式版 同城方案

2AZ



- 单分片支持10T+数据
- 单分片性能为HA版/开源MySQL 7倍

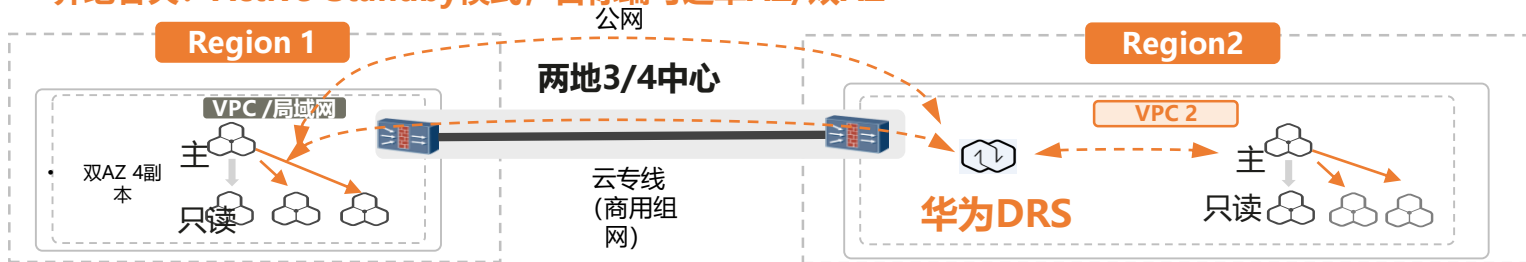
3AZ



- 单分片支持10T+数据
- 单分片性能为HA版/开源MySQL 7倍
- 整体性价比最高

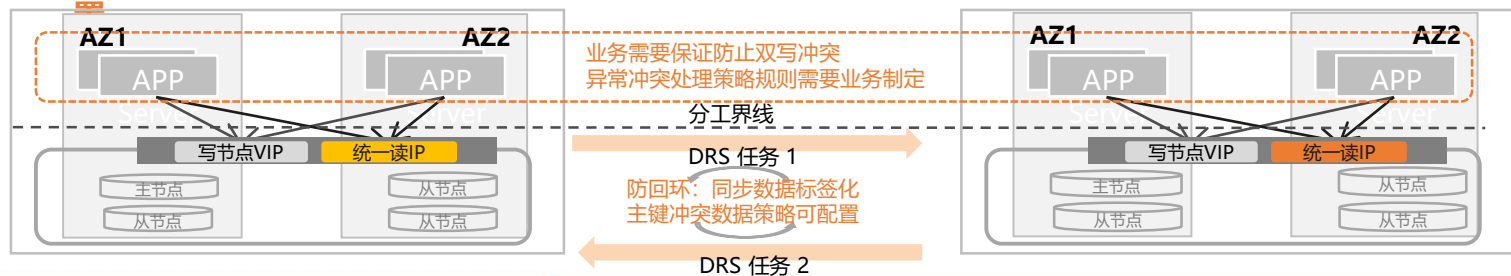
GaussDB for MySQL 分布式版 异地容灾/双活

- 异地容灾：Active-Standby模式，目标端可选单AZ/双AZ



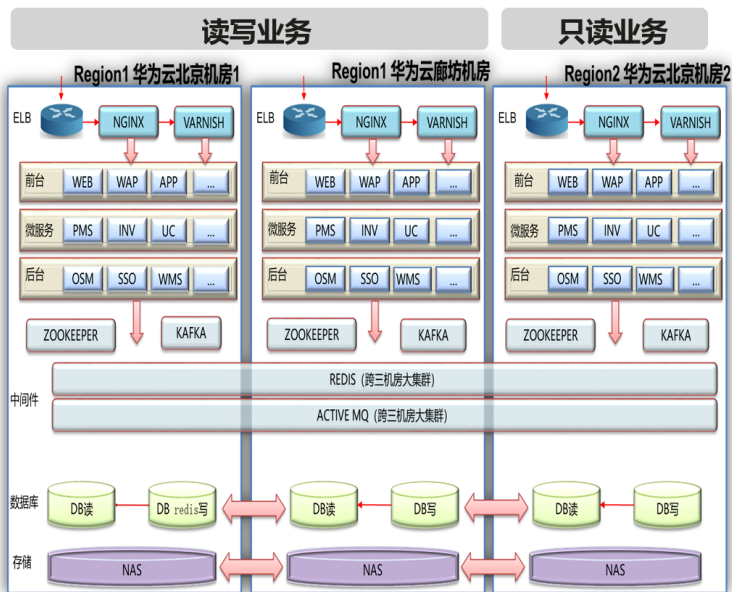
- ☐ **秒级RPO&RTO**: 优于《信息安全技术信息系统灾难恢复规范》灾备等级第五级，物理距离1500公里，秒级时延（网络良好、小事务时）
- ☐ **自由主备切换**: Active-Standby 模式，可在业务数据库和灾备数据之间自由升主降备，填补业界常见的单向灾备短板。
- ☐ **数字化管理**: 清晰展示了RPO、RTO、带宽、网络、读取、写入、数据健康报表等，提供最清晰可控的灾备环境。

- 异地双活：消费者云异地容灾Active-Active模式，解决回环问题，提供冲突策略配



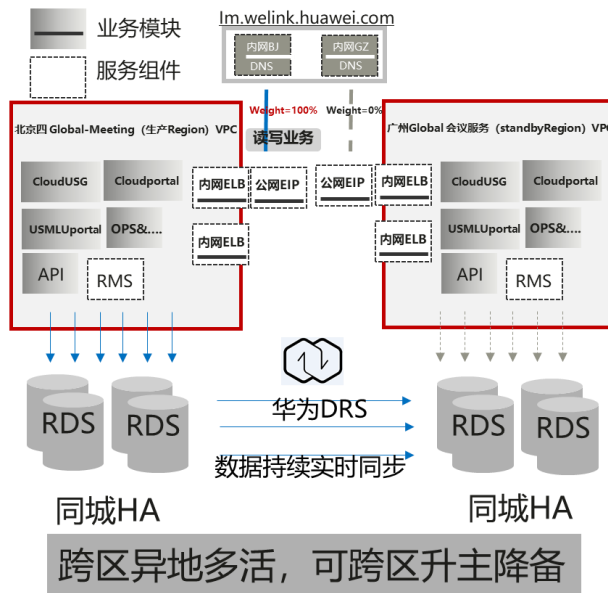
VMALL 华为商城

华为商城（VMALL）是华为公司旗下的自营电子商务平台，是以营造用户的移动信息生活为服务宗旨的互联网商务平台。通过华为云提供的服务，实现终端的销售和节日促销/新机发售的支撑。

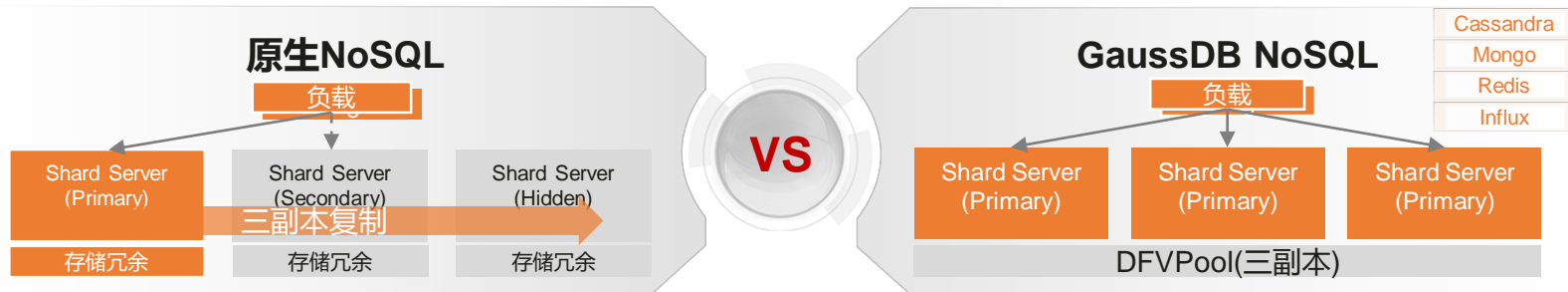


华为云WeLink

华为云WeLink，是华为18万员工的手机协同办公平台，融合消息、会议、邮件、知识、能力开放等，打造企业数字化办公协作平台，实现团队、知识、业务、设备的全面连接。



GaussDB NoSQL: 国内领先的多模NoSQL服务



原生主备技术: RPO可靠性较低、节点扩容时间长

独创的ALL-PRIMARY技术: 多写, RPO=0, 分钟级节点扩容

故障切换时间: 快20倍

节点扩容时间: 快20倍

性能: 优1.5倍

数据备份时间: 快10倍

备份恢复时间: 快7倍

JAC 江淮汽车
—— 实·现未来 ——

天地图
MAPWORLD

GAME

锦江 WeHotel
全球酒店共享服务平台



架构革新 11th
高效可控

IT168.com

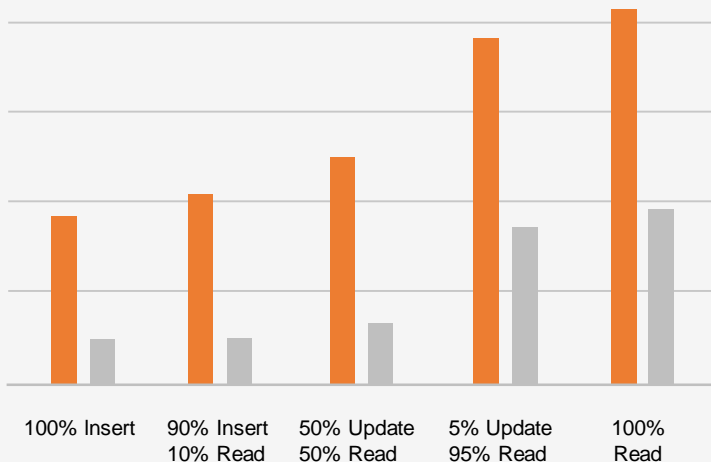
ChinaUnix

ITPUB

大企业案例： GaussDB NoSQL， 相对于开源自建和基于开源服务化的优势

江淮汽车：极致性能

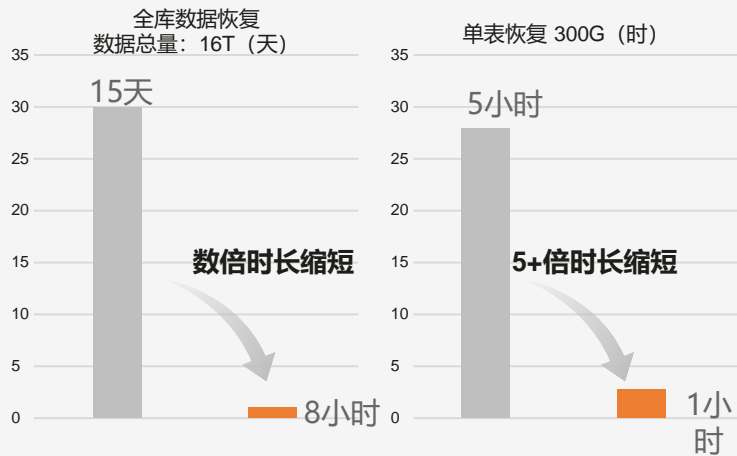
■ GaussDB NoSQL ■ 自建



JAC 江淮汽车
—— 实·现未来 ——

天地图：快速备份恢复

■ GaussDB NoSQL ■ 自建



天地图
MAPWORLD

THANKS

