



# 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GaussDB加速企业数字变革 与实践及新品发布

张昆 | 华为云数据库产品总监



架构革新 11<sup>th</sup>  
高效可控



# 目录

## CONTENTS



1

华为云数据库发展及GaussDB全新升级

2

GaussDB产品革新与技术演进

3

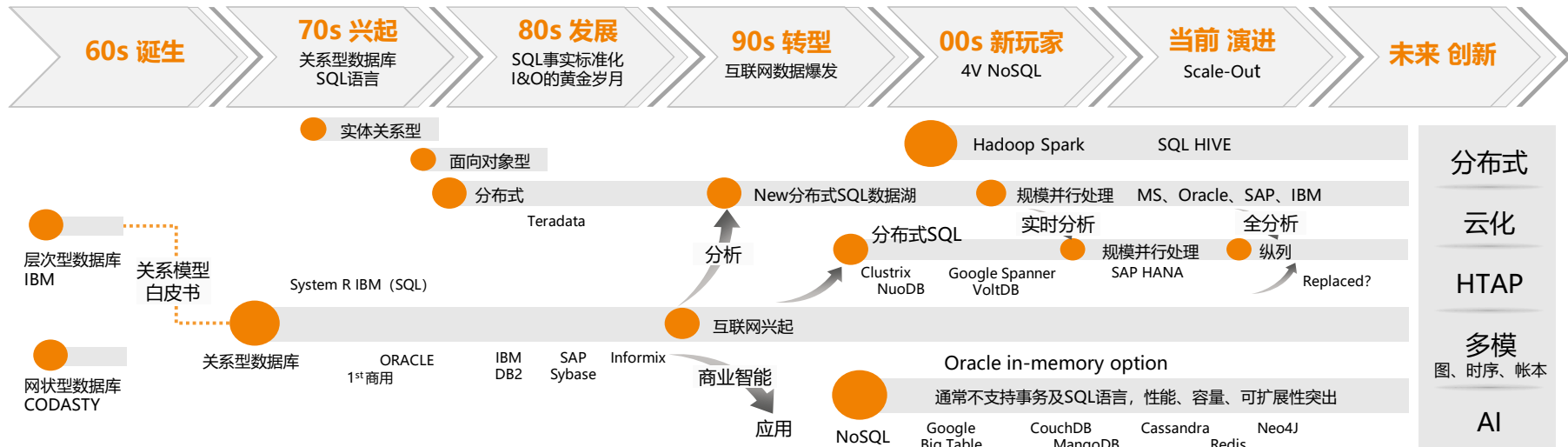
GaussDB客户实践：加速政企数据库转型

4

华为云数据库新品发布



# 数据库历史回顾和趋势：云化，融合是方向



## 不同类型的数据库基于不同的理论基础，满足不同的数据管理、处理需求

	关系型数据库 Oracle/MySvQL	键值数据库 Redis	宽列数据库 Cassandra	文档数据库 MongoDB	图数据库 Neo4J	ML数据解决方案 Spark/Tensorflow	数据湖解决方案 融合数据分析
应用场景	事务处理	储存用户信息等	日志等	Web应用等	图分析、搜索	数据分析、特征抽取	全场景
数据模型	关系表	K-V键值对	以列簇式存储	KV键值对, V结构化	图模型	稀疏矩阵	关联数组
数学基础	关系代数, 集合论	hash table	KV二维数组	嵌套KV	图论	线性代数	关联代数

# 华为数据库依托华为云整体解决方案和全栈软硬件创新能力， 为客户服务

数字化政府 智慧城市 智能制造 智慧金融 行业互联网 智慧园区 智慧零售 智能网联汽车 云游戏



基础服务

AI企业智能

数据库服务

开发者

安全

IoT物联网

企业应用

专业服务

18 类 200+ 服务, 190+ 解决方案, 10000+ 商业合作伙伴



混合云解决方案

统一应用生态

统一高阶服务

统一运维管理



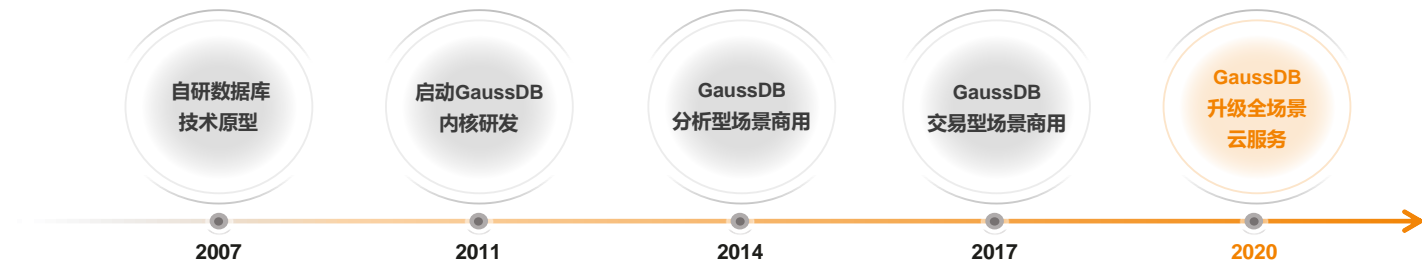
多元算力



# 华为囊括全球1000+数据库专业人才，战略投入超13年



7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（套）



# GaussDB数据库：为政企客户打造传统商业数据库一站式上云解决方案

## 选型安心

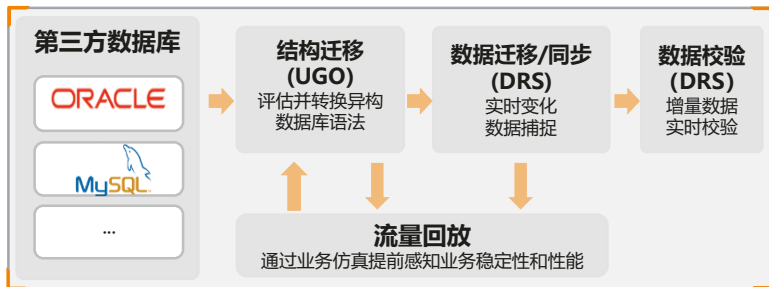
全场景业务，全开放生态，安全可靠



单机社区版本开源网站: <http://opengauss.org>

## 迁移放心

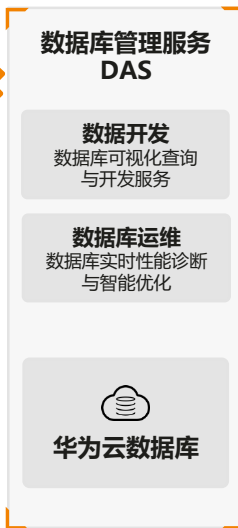
数据库结构+应用+数据一站式迁移



不丢数据，不中断/影响业务，迁移成本低，提前仿真感知，透明可视化

## 管理省心

无缝协同上层应用





# 目录

CONTENTS



1

华为云数据库发展及GaussDB全新升级

2

GaussDB产品革新与技术演进

3

GaussDB客户实践：加速政企数据库转型

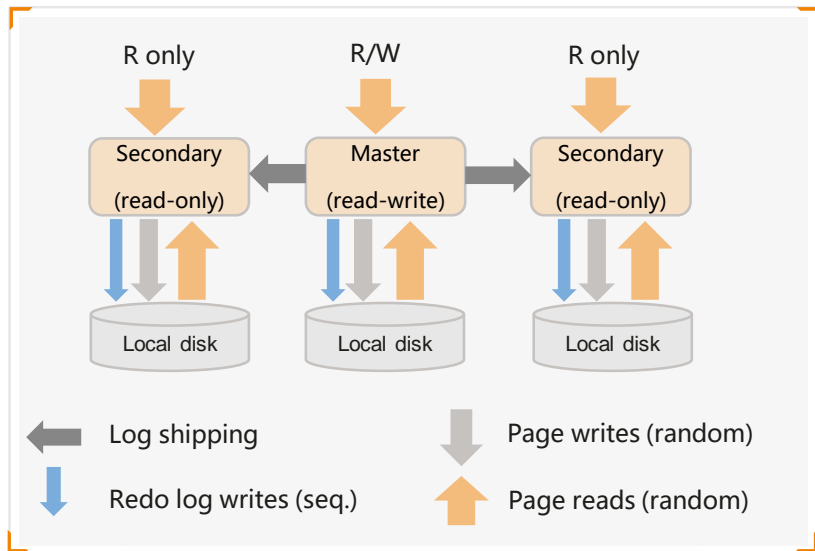
4

华为云数据库新品发布



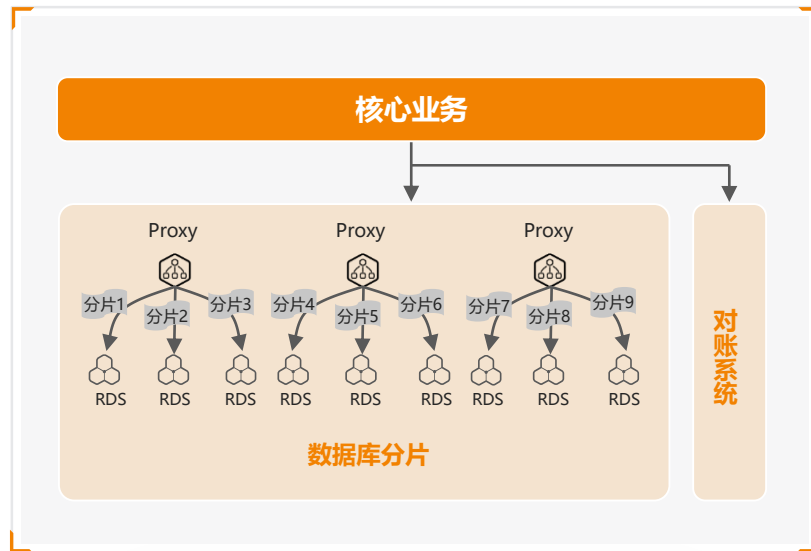


## 新场景新挑战：在政企、金融等企业级客户业务场景更为复杂



开源架构已经30年，传统HA方案：

- 半同步模式在极端场景可能丢数据
- 复制会导致性能下降



传统解决方案创新，但依旧存在问题：

- 对账系统解决高可用/高可靠问题 → 事后业务成本高
- 核心数据库分片，应对业务演进 → 有一致性问题

## 来自市场的核心诉求

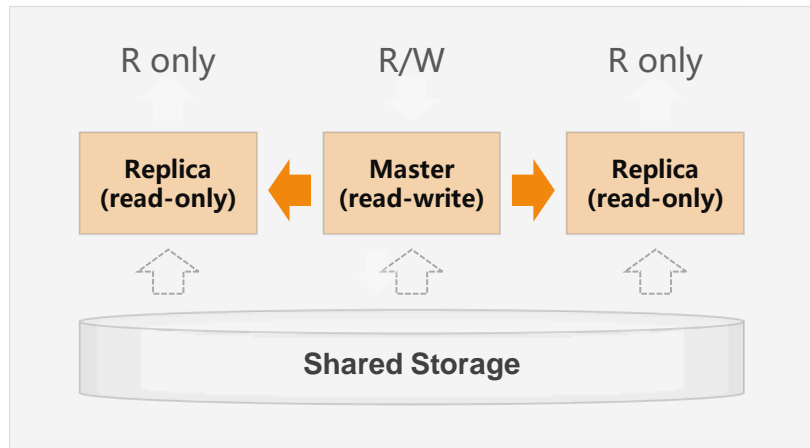
TopN 需求和痛点	需求描述
海量数据存储	支持大数据量业务
分布式高扩展	自动化分库分表或者非分库分表，应用透明
无缝弹性扩容	分钟级弹性扩容，应用透明
强一致事务	支持分布式事务的强一致性，数据0丢失
高可用	支持跨AZ高可用、跨 Region 容灾
高并发性能	支持大并发下的极致高性能
非中间件式架构	非 DDM 类方案（或者非 DRDS）



# GaussDB(for MySQL)全自研企业级OLTP数据库

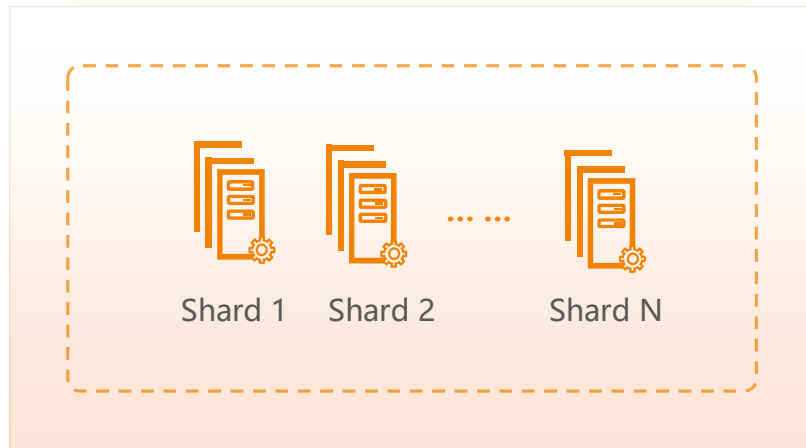
## 主备部署

最大支持1写15读，分钟级添加只读实例，支持读写分离，自动负载均衡



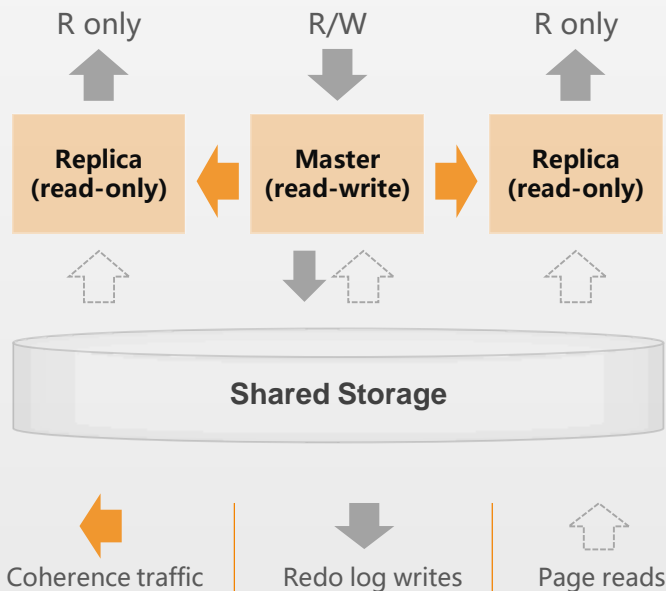
## 全分布式部署

分布式高扩展：数据按shard划分，读写负载线性扩展，满足大规模业务量场景



华为云底座

# GaussDB(for MySQL)主备版架构和优势



## 极致可靠

数据0丢失，故障闪恢复，支持跨AZ 高可用。



## 多维扩展

计算节点双向扩展，横向扩展：支持1写15读横向扩展，纵向扩展：在线弹性扩容，按需计费



## 海量存储

单实例扩容数据达128T，无需分库分表，超低应用改造，业务急速上云



## 创新自研

Cloud-native分布式数据库架构，基于华为最新一代DFV 存储计算存储分离，保证扩展性价比；数据库逻辑下推存储，最小网络负载，极致性能



## 卓越性能

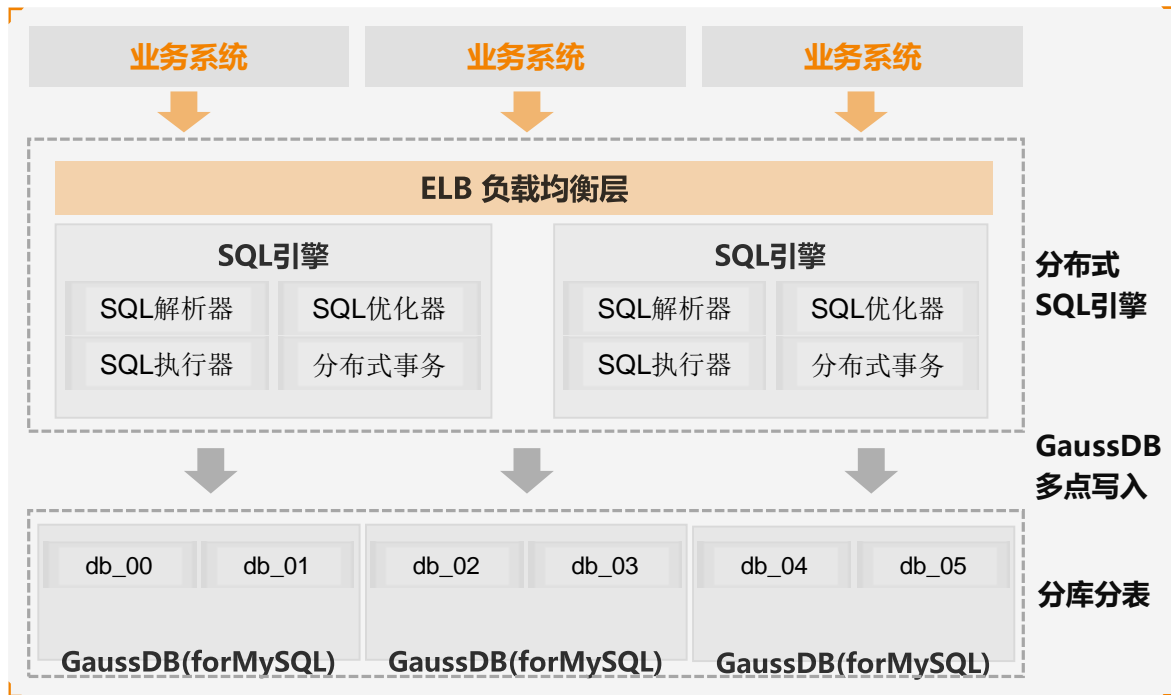
性能最高提升至原生MySQL 的7倍，100%兼容 MySQL，业界领先



## 尖端硬件

基于V5CPU+Optane DC SSD+RDMA 网络行业领先硬件组合，数据处理既稳又快

# GaussDB(for MySQL) 分布式架构和优势



DFV 分布式存储层

## 兼容开源MySQL生态



- 语法
- 工具
- 连接器
- 开发生态

## 超高性能，满足海量数据高并发场景



- PB级数据
- 百万级TPS
- 千万级QPS

## 极致架构成就极速扩展

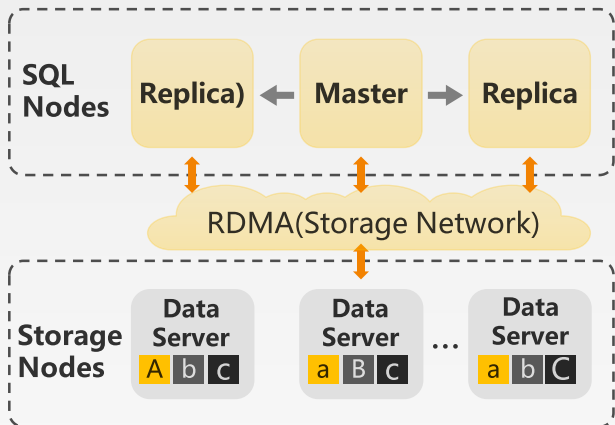


- 300s分片扩展
- 自动扩展存储空间
- 秒级热点数据迁移

# GaussDB(for MySQL)能力和演进性优势明显

## 竞品

计算存储分离



VS

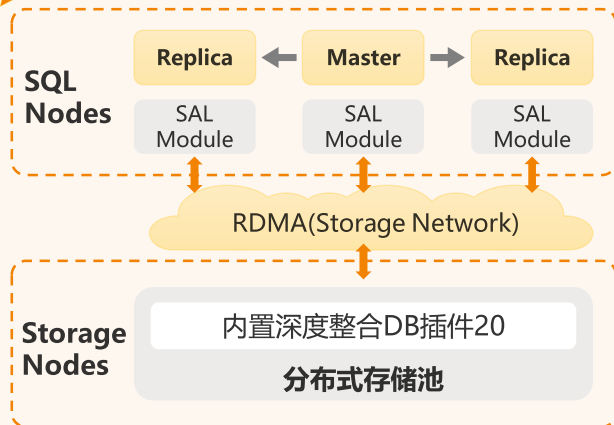
性能更强  
提升30%

可靠性更佳  
支持跨AZ

架构更优

## GaussDB(for MySQL)

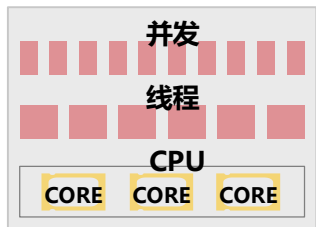
计算存储分离 + 数据库逻辑下推



# GaussDB(openGauss)数据库，灵活应对核心交易系统和海量事务型场景

## GaussDB(openGauss)主备版 聚焦企业核心交易场景

GaussDB(openGauss)  
支持Kunpeng的企业级数据库



极致性能、极致高可用

开源社区版：<https://opengauss.org/zh/>

## GaussDB(openGauss)分布式版 聚焦海量事务扩展场景



政府



交通



金融



能源



运营商



1GB/天

一个互联网用户



10GB/天

一个智慧家居



64TB/天

一辆自动驾驶汽车



200TB/天

一架联网飞机



1PB/天

一家数字化互联工厂



50PB/天

一个大型新闻网站

### 数据新基建

算力 | 存储 | **数据库** | 大数据 | AI

异构计算框架：端到端多样化算力



# GaussDB(DWS), 全场景企业级数据仓库

## 产品定位

新一代、全场景、云数据仓库

### 统一内核 统一架构

- 一套内核、一套架构支撑公有云、混合云，用户体验一致
- 同时支持虚拟机(ECS)、裸金属(BMS)、物理机(HCS)模式

### 关键 竞争力

#### 高性能

TPC-DS性能业界领先

#### 高扩展:

2048节点, >10PB

#### 高可用

**强一致性:** 集群内 RPO=0,  
RTO<30s; 在线扩容

#### 企业级特性:

逻辑集群、SQL on Anywhere、  
融合分析(+AI, +IoT)、GDPR

## 主打场景

数据仓库、数据集市、自助分析

### 超大规模 金融数仓

- 某国有大行单集群**242节点四路全闪存服务器**, 装机容量**10PB**
- 完成权威机构**2048节点**测试
- 连续3年入选Gartner, **专利180+**

### 典型行业

#### 金融

银行、保险、证券.....

#### 政府

财政、税务、公安.....

#### 能源

电力、石油.....

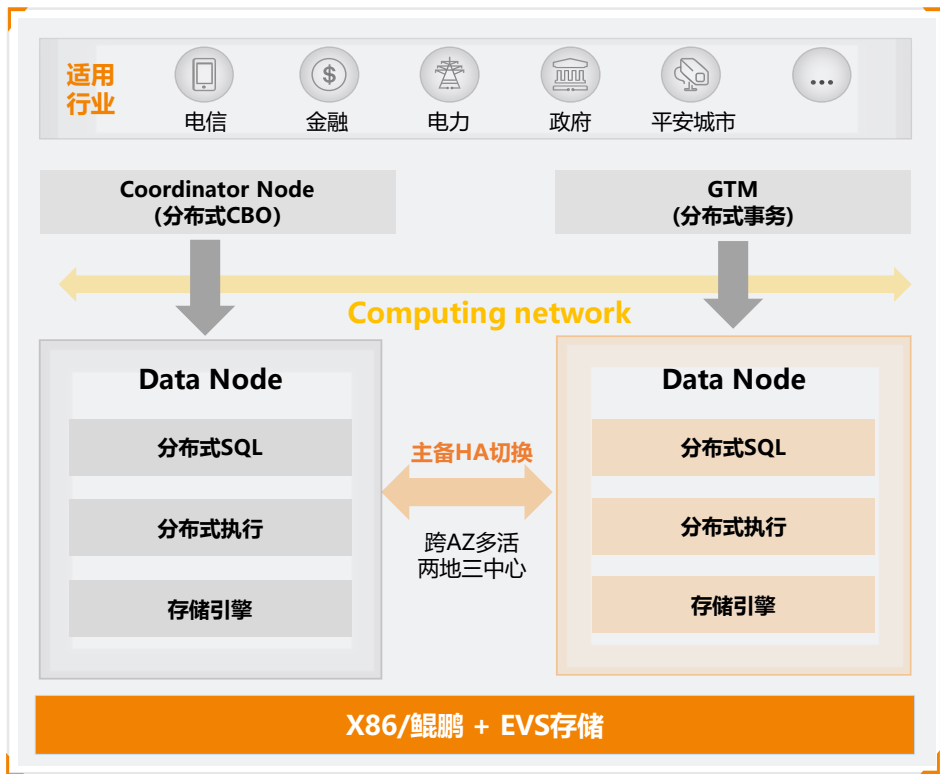
#### 交通

地铁、机场、高速公路.....

#### 酒店、商超

#### 物流

# GaussDB(openGauss)总体架构和优势



## 工具集



异构迁移  
UGO



数据同步  
DRS



运维管理  
DAS

## 高性能: 支持高吞吐强一致性事务能力

鲲鹏2路服务器

- 分布式强一致: 32节点1200万tpmC

## 高可用: 双活和两地三中心高可用

- 集群内HA, 数据不丢失, 业务秒级中断
- 同城跨AZ容灾, 数据不丢失, 分钟级恢复
- 两地三中心部署

## 高扩展: 容量和性能按需水平扩展

- 256节点扩展能力, 卓越线性比
- 在线扩容

## 易管理: 易迁移, 易监控, 运维

- 兼容SQL2003标准语法+企业扩展包
- 数据复制、监控运维、开发工具

# 面向企业新型应用的云原生多模GaussDB NoSQL系列服务



极致性能

3倍

读写性能提升

自动弹性伸缩

Auto Scaling



企业级可靠性

N-1个

节点故障容忍

10倍

故障重构性能

10倍

备份恢复性能



灵活全托管

分钟级

计算资源扩容

秒级

存储资源扩容

100T

海量存储

文档型接口  
兼容MongoDB

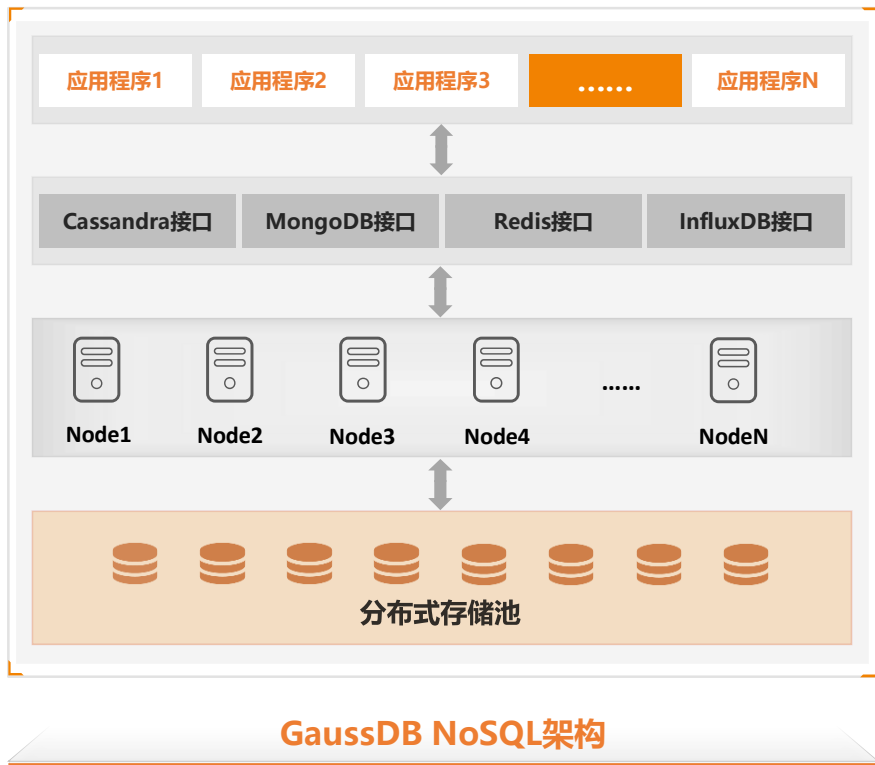
Key-Value接口  
兼容Redis

宽列接口  
兼容Cassandra

时间序列型接口  
兼容InfluxDB

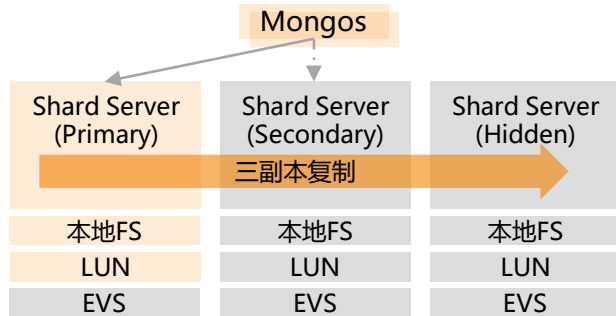
多模  
GaussDB  
NoSQL  
系列服务

# 计算存储分离架构，极致性价比

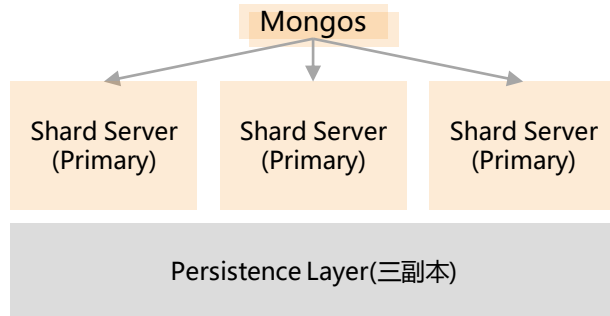


# GaussDB NoSQL系列服务计算存储分离的架构创新

## 社区版MongoDB



## GaussDB(for Mongo)



	社区版MongoDB	GaussDB(for Mongo)
性能	一个Shard Server节点承担读写压力	三个Shard Server分担读写压力，相同成本2~3倍性能。
成本	存储成本：两层复制，3*3共9副本存储； 存储计算绑定，同步扩展； 起配成本：副本集3节点，集群模式11节点；	存储成本：仅存储层3副本复制 存储计算分别按需扩展，降低20%~40%成本； 起配成本：只支持集群模式，7节点起配。
HA	副本集主备切换，10s。	ShardServer间Failover，<5s。
扩容	小时~天级；扩容步长3节点。	分钟级完成，100倍以上性能提升；扩容步长1节点。
备份恢复	逻辑备份，小时级。	基于快照的物理备份，分钟级，10倍性能提升。
生态兼容	100%兼容，屏蔽高危、管理类命令。	屏蔽convertToCapped、已有表重新Shard；

## 主要技术创新

### 存储计算分离

- 存储计算分别按需扩展，降低成本；
- 基于共享存储，Rebalance不迁移数据；
- 3AZ容灾；

### 复制集卸载到分布式存储

- 减少存储副本数量；
- 所有ShardServer均可处理业务；
- 分布式存储基于分片复制，更好的聚合IO性能和故障重构性能；

### RocksDB存储引擎

- 基于LSMTree，相比WT，写性能更好；
- 本地SSD读Cache，优化读性能；

### 基于快照的物理备份

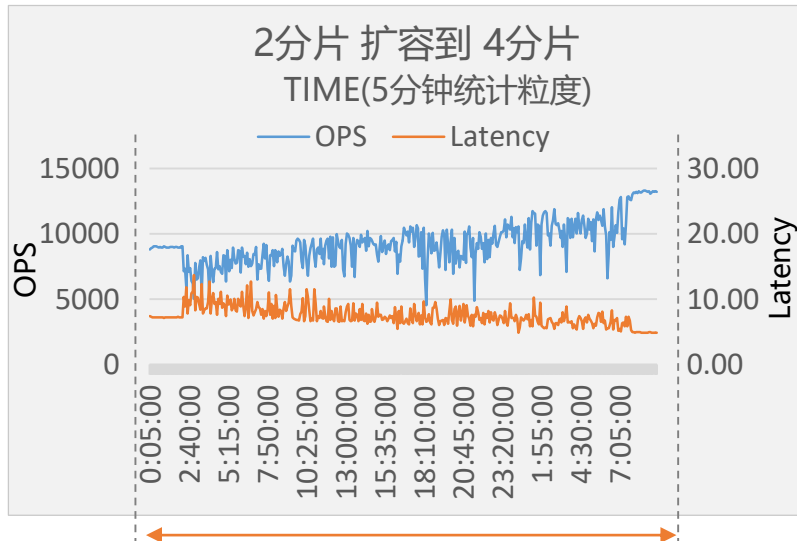
- 避免逻辑备份dump数据，性能更好；
- 明确的备份时间点；

### On going

- 性能优化：基础设施、线程池、存储RDMA；
- Auto Scaling：自动根据业务负载扩容集群规模，降低用户50%以上成本；
- 瞬时恢复、增量备份、表级备份、任意时间点恢复；

## 实现分钟级节点扩缩容

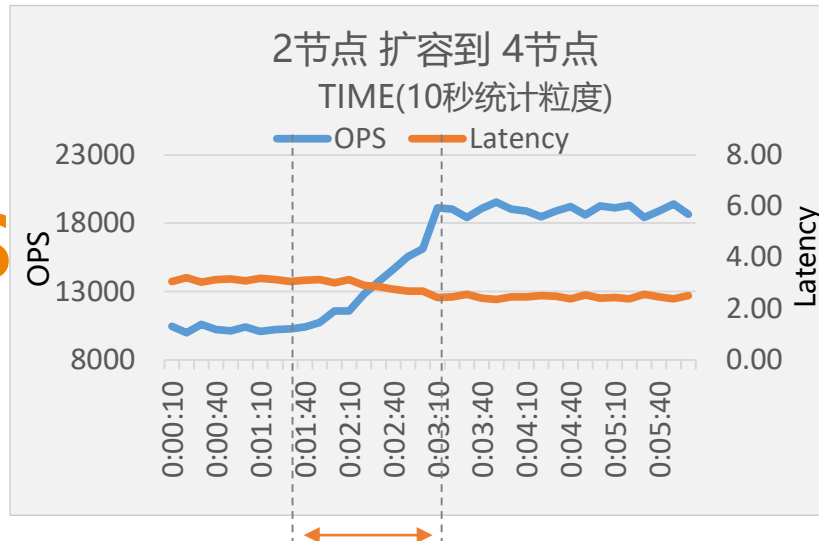
### 社区自建扩容迁移



扩容耗时 十小时级

VS

### 华为云GaussDB(for Mongo)



扩容耗时 分钟级

# 未来，华为云数据库将持续技术创新与服务好客户





# 目录

## CONTENTS



1

华为云数据库发展及GaussDB全新升级

2

GaussDB产品革新与技术演进

3

GaussDB客户实践：加速政企数据库转型

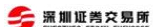
4

华为云数据库新品发布



# 在500+大客户规模商用，遍布金融、政府、电信、能源、交通、物流、电商等行业

## 泛金融



## 泛政府

### 22+省级：

北京、重庆、上海、江苏、广东、新疆、贵州、...

### 100+地级市：

深圳、广州、厦门、西安、无锡、苏州、青岛、南京、中山、宜昌、天津、石家庄、扬州、连云港...

涉及财政、税务、医保、自然资源...

## 能源&交通



成都空港城市发展集团

## 酒店、互联网 快递、娱乐



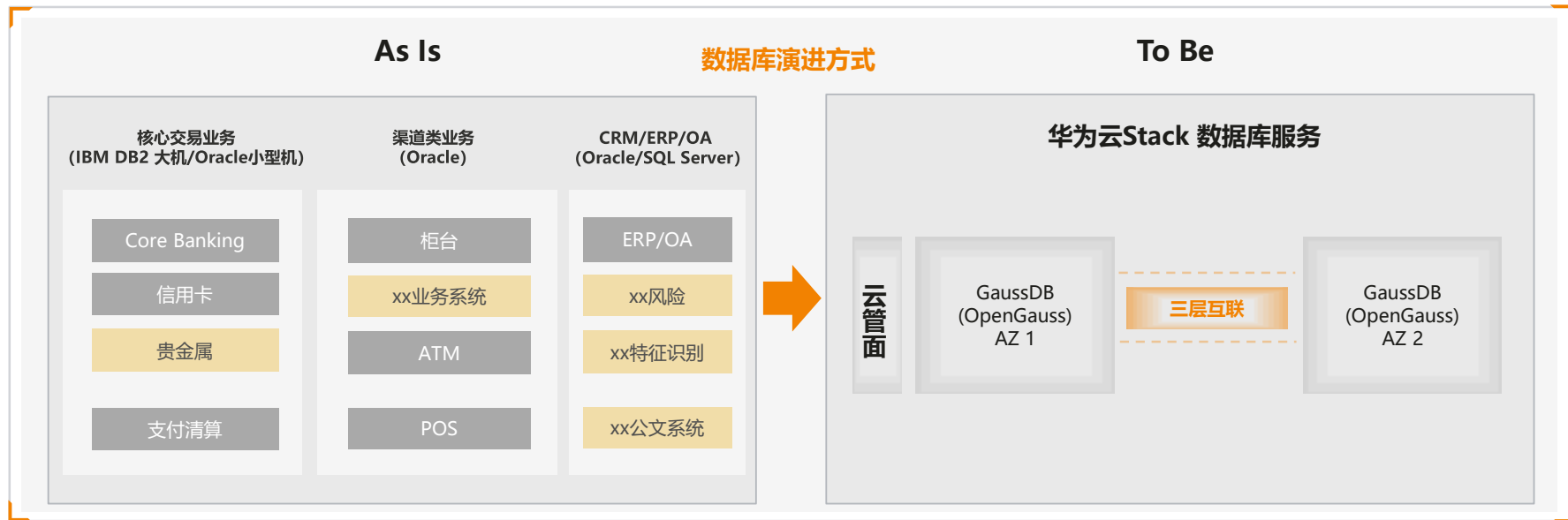
## 电信运营商 软件



## 制造、电商

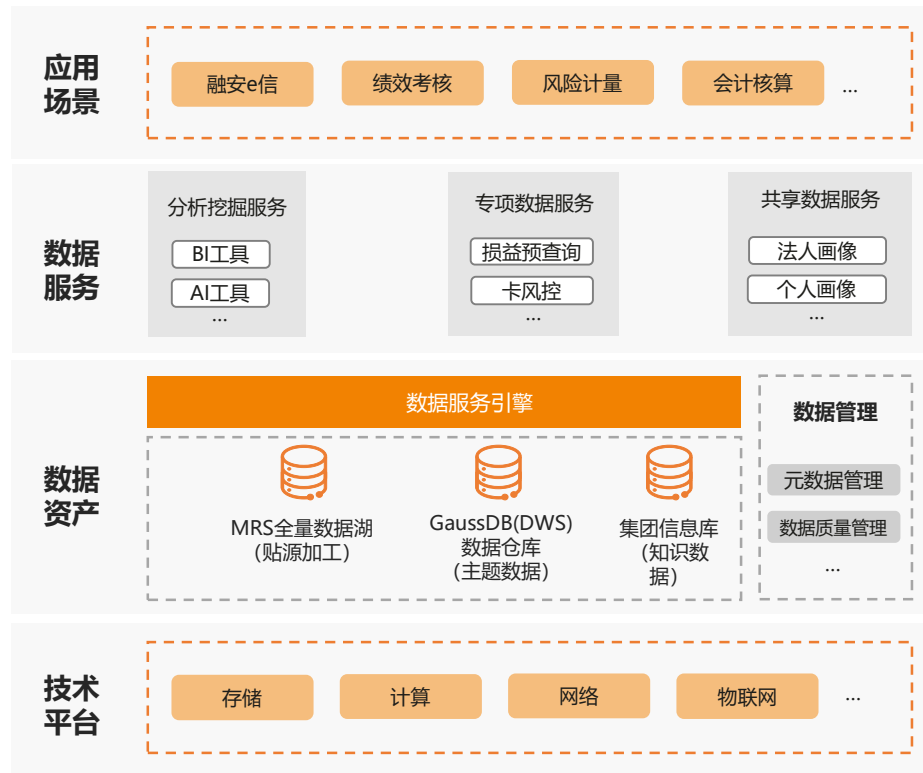


# 国有某大行核心业务已商用GaussDB(openGauss)



- **高可用**：同城单Region多AZ互联支持同城双活，金融级支持三层组网的管控高可用部署方案
- **性能线性扩展**：支持集群水平扩展，基本联机性能与O持平
- **弹性部署**：部署实例及应用要快速响应业务需求，数据库支持容器化部署

# GaussDB(DWS)助力国有大行构建超大金融数仓



## 客户痛点

- 海量数据达45PB，资源管道开发周期长，导致需求大量积压
- 烟囱式系统建设，形成数据孤岛，各部门间存在数据壁垒

## 解决方案

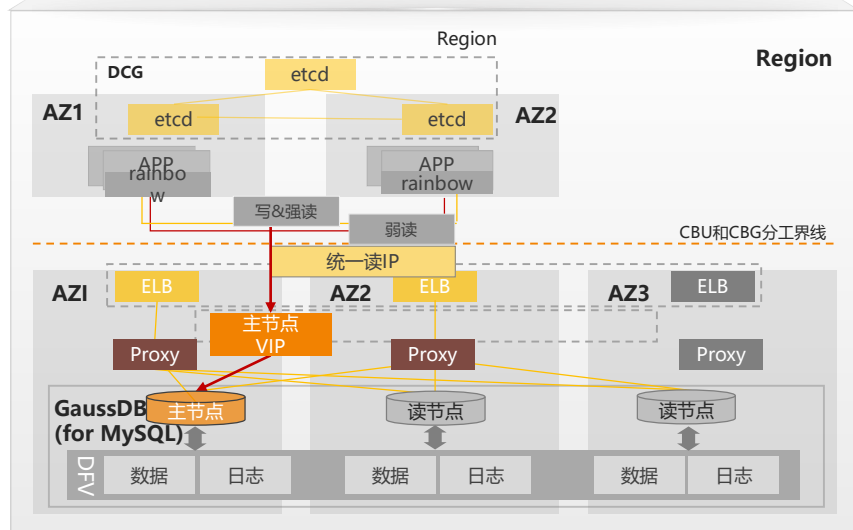
- 采用GaussDB(DWS)替换Teradata、exadata,构建核心数仓和数据集市，通过企业级多租户和资源隔离能力支持批量处理和联机业务的混合负载
- 通过MRS建立流处理平台提供半结构、非结构化数据实时计算服务

## 客户价值

- 业务天级变更**，新业务数据分析平台三天内上线，数据分析结果实时展现
- 支持上万分析师同时在线自助开发**，满足客户开发需求，支持业务开发效率
- 支撑行内行外业务，对外提供风险信息服务，服务200+同业，3W+企业客户，**每年带来x亿收入**

# 华为消费者业务数据库全面切换上云，3AZ数据0丢失

## 华为消费者业务数据库全面切换上云



### 核心诉求

跨AZ RPO为0的前提下，RTO < 10s

性能相较单AZ场景损耗 < 10%

### 关键技术

日志强一致

页面Gossip保障可靠性和性能

# GaussDB助力开天创世游戏火爆开服

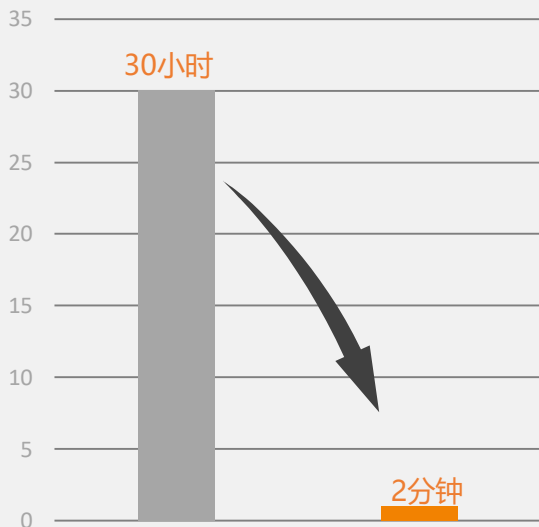
急速无感扩容，百倍超越友商

开天创世 GAME

游戏场景：急速扩容

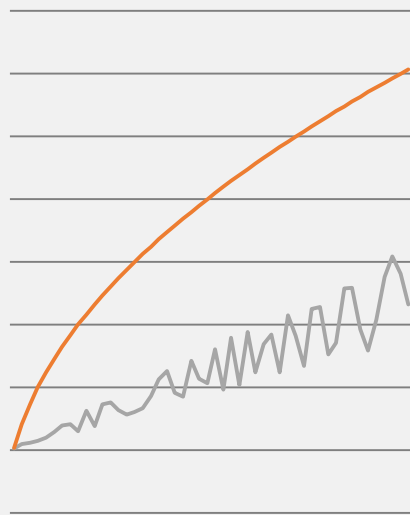
- 游戏开服6小时内需多次扩容
- 表级时间点恢复支持游戏快速回档
- 扩容期间性能稳定，不影响游戏体验

扩容效率百倍提升



■ 基于ECS自建或基于开源服务化的方案

急速扩容，性能平稳



■ 华为云多模GaussDB NoSQL系列服务

# 目录

## CONTENTS



1

华为云数据库发展及GaussDB全新升级

2

GaussDB产品革新与技术演进

3

GaussDB客户实践：加速政企数据库转型

4

华为云数据库新品发布





# 传统数据库迁移的痛点与诉求

## 传统数据库迁移面临的问题

1

源库的运行状态如何，容量和复杂度如何，库表之间关系如何？

2

如何选择合适的目标库及其规格，改造风险和工作量有多大？

3

从源库迁移到目标库，有哪些对象不兼容，在目标库是否有替代方案？

4

对于应用程序改造，如何动态或者静态抓取到应用改造的SQL？

5

人力进行语法改造，成本巨大，如何做到迁移过程半自动化或者自动化？

6

改造过程涉及业务重构时，如何为用户提供合适的改造方案？

7

源库对象较多，如何保障源库对象顺利迁移到目标库？

8

如何保证结构迁移和数据迁移时，不影响源库/目标库正在运行的业务？

9

源端按业务模块迁移时，源和目标如何进行持续的增量数据同步？

10

语法改造后，如何保证性能符合预期，如何对改造后的对象进行优化？

11

迁移完成时，如何保障迁移结果正确，包括结构的准确性与数据的准确性？

12

迁移完成后，如何保障云上数据库的开发运维与性能的持续优化？

评估

改造

迁移

优化

# 新品发布：数据库和应用迁移 UGO

数据库和应用迁移 UGO (Database and Application Migration UGO, 简称为UGO) , 是专注于**异构数据库对象迁移**和**应用迁移**的专业化工具。通过**预迁移评估**、**结构迁移**二大核心功能, 实现主流商用数据库到华为云数据库的自动化搬迁, 助力用户**轻松实现一键上云、一键切换数据库**的目的

## 支持场景

### 源端

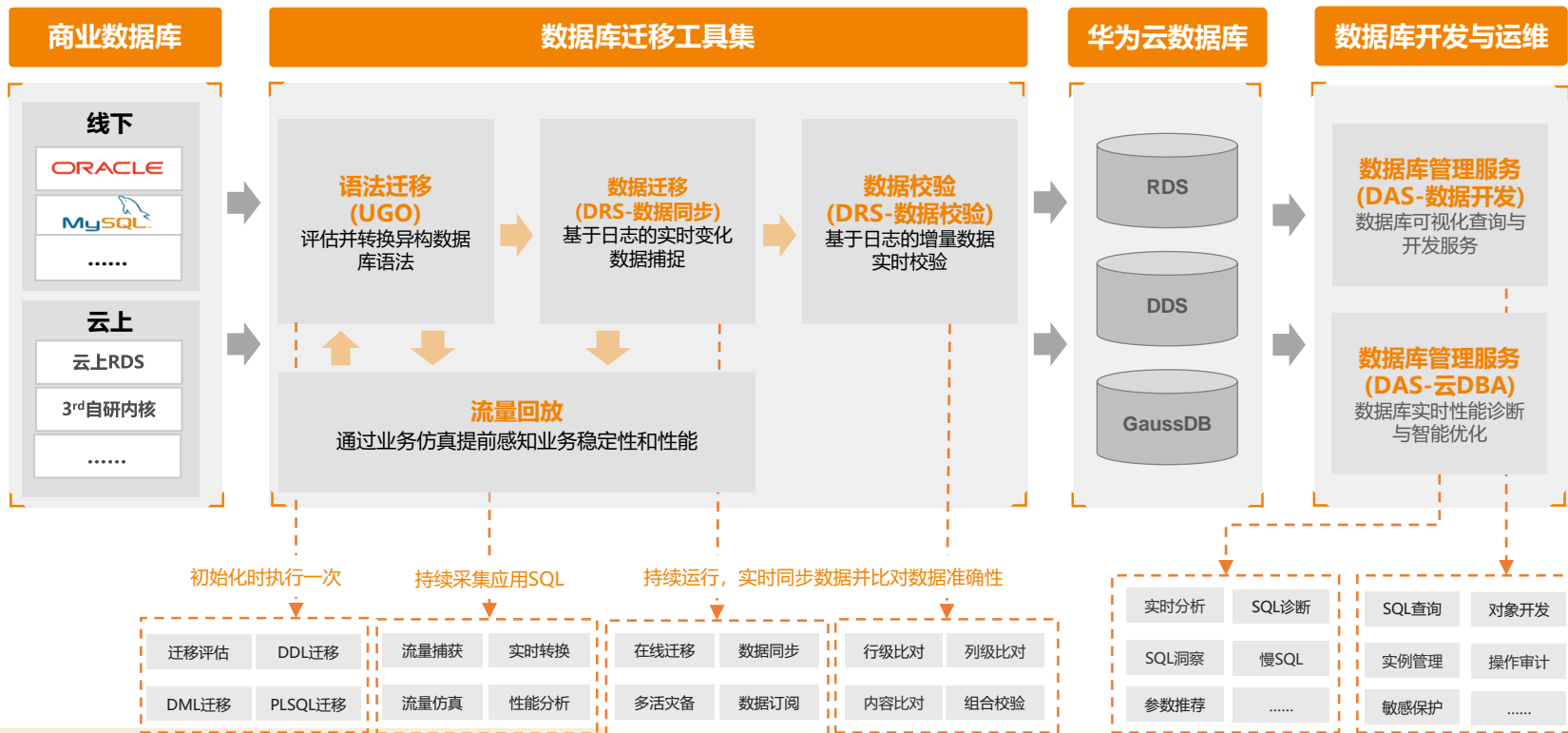
- Oracle 10g
- Oracle 11g
- Oracle 12c

UGO

### 目标端

- RDS for MySQL
- RDS for PostgreSQL
- GaussDB(for MySQL)
- GaussDB(for PostgreSQL)
- GaussDB(openGauss)

# 新品发布：金融行业核心数据库上云解决方案



# THANKS

