



20000节点数仓集群 在大型商业银行的落地实践

建信金科

基础技术中心 陈晓新

目录

01 基于龙跃MPP的超大规模数据仓库建设

02 龙跃MPP技术架构和带来的应用革新

03 基于龙跃MPP的大数据未来技术展望

龙跃MPP DB

龙跃MPP DB——新一代云原生数据仓库产品



建信金科
CCB Fintech

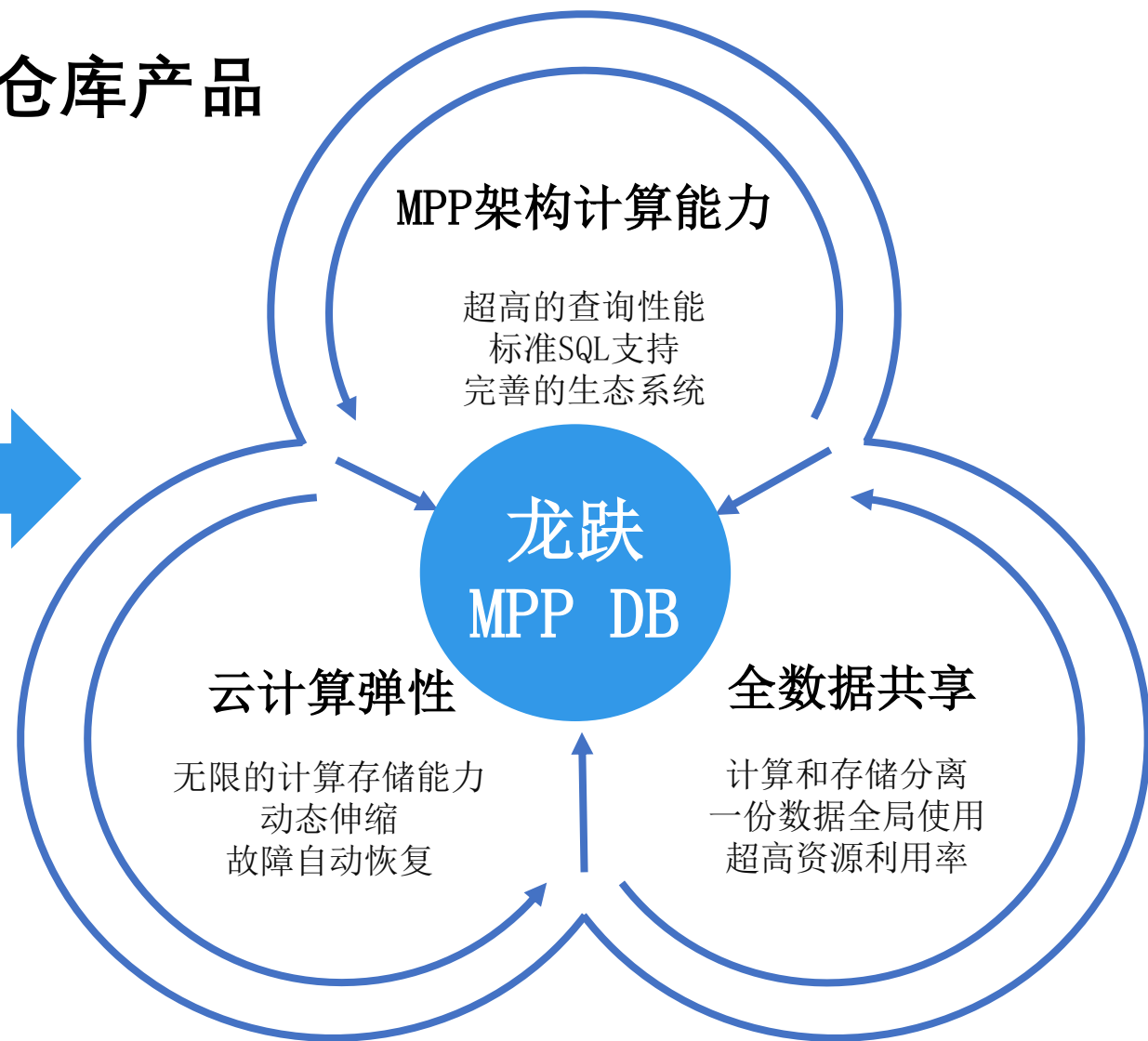


HashData

金山云

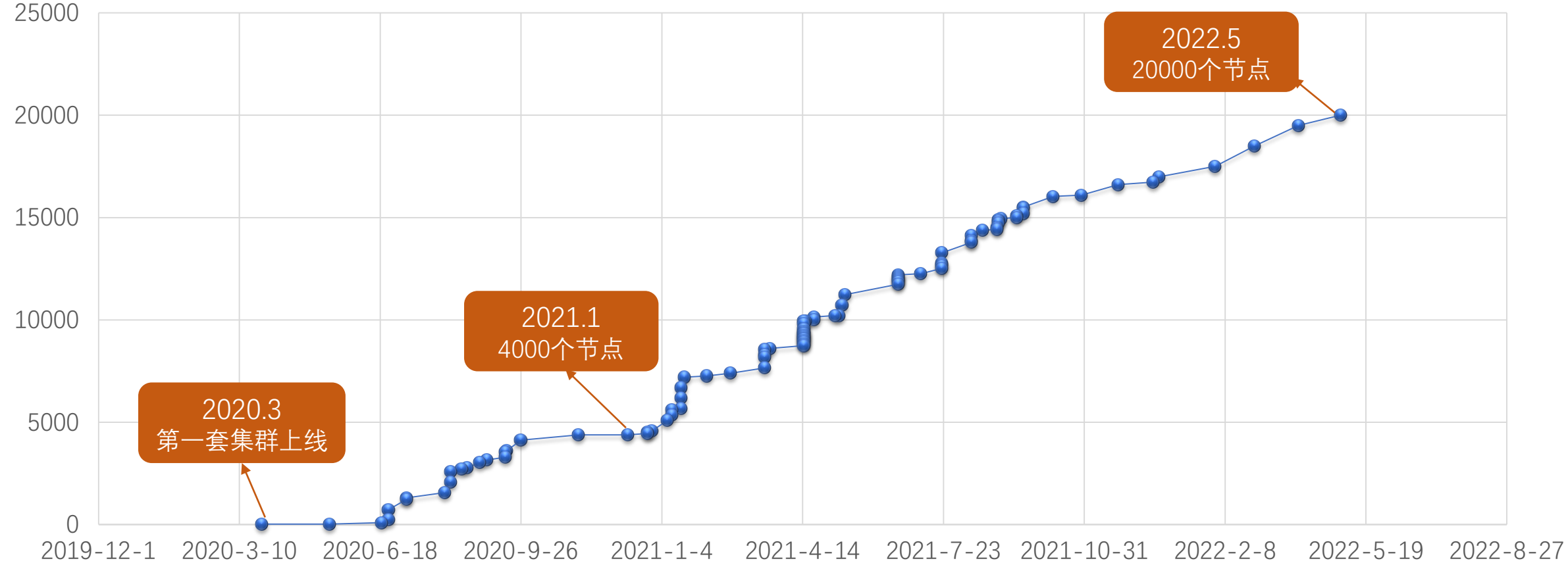


其他数据库、存储等合作公司



基于龙跃MPP DB的超大规模数据仓库建设

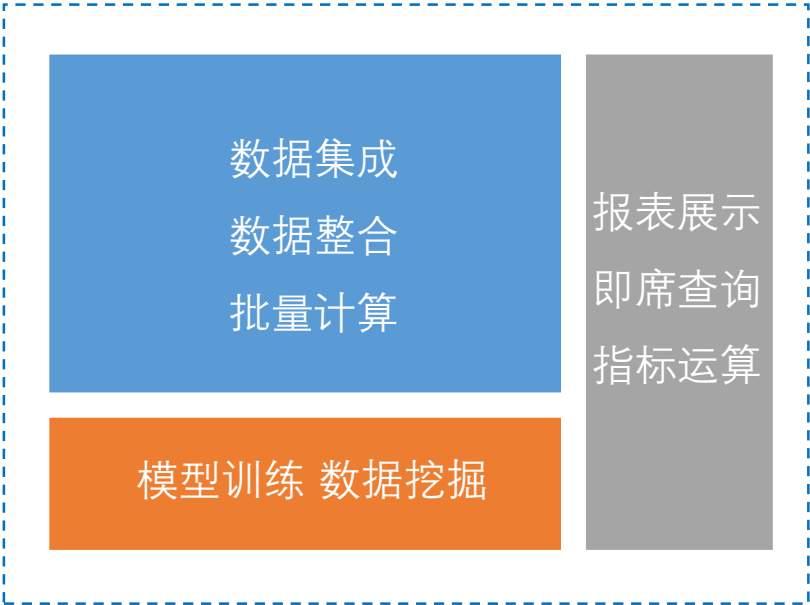
龙跃MPP DB 运行现状	集群规模	数据量	表数量/对象数	负载情况
	20000+节点	18PB	百万/千万	每天运行作业数达到百万级别，SQL数千万级别



基于龙跃MPP DB的超大规模数据仓库建设



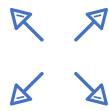
基于龙跃MPP DB的超大规模数据仓库建设



龙跃MPP DB



计算存储分离，
数据充分共享



百PB级数据，并
发能力线性扩展



金融级高可用
准0-RPO容灾



动态扩容、升
级，智能运维

基于龙跃MPP DB的超大规模数据仓库建设

恒丰银行

9套计算集群
80+ 台物理机
规划数据量500TB

北京银保监

3套计算集群
1+64 ; 1+64 ; 1+128
存储数据量150TB+

进出口银行

24台物理机
规划数据量100TB

国家开发银行

24台物理机
规划数据量100TB

云南政务

河北政务

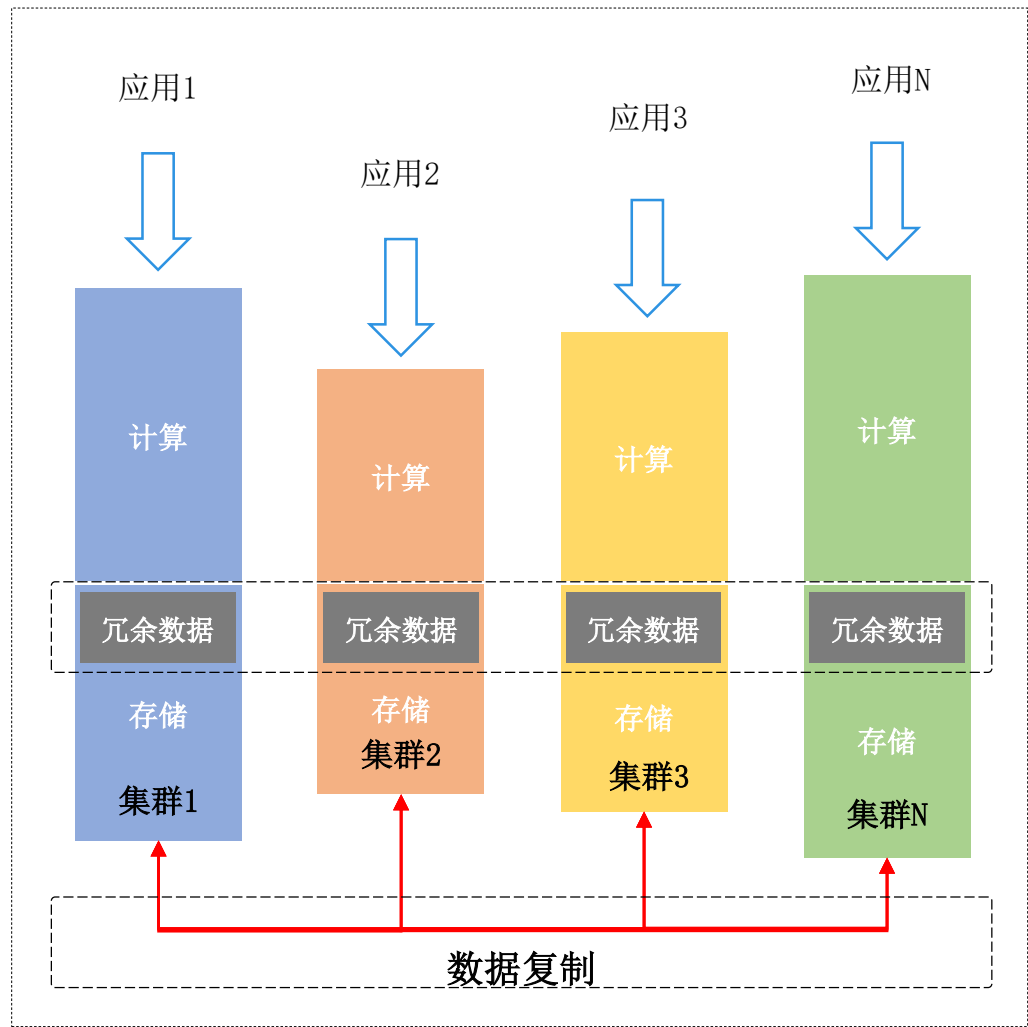
目录

01 基于龙跃MPP的超大规模数据仓库建设

02 龙跃MPP技术架构和带来的应用革新

03 基于龙跃MPP的大数据未来技术展望

基于传统MPP的数据仓库建设问题



运维困难

- 传统物理机/虚拟机架构，运维管控难度大

数据孤岛

- 应用烟囱式设计，一个应用搭建一个独立集群，导致严重数据孤岛

数据冗余

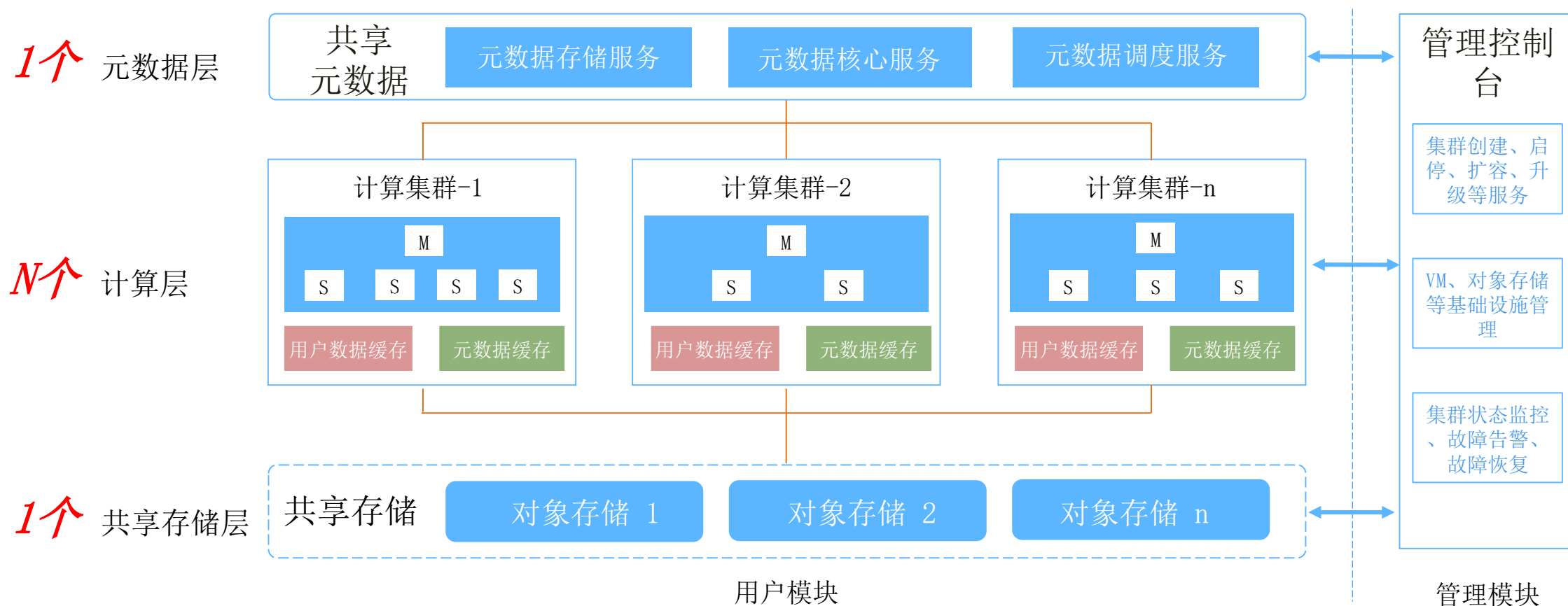
- 每个集群都存储大量冗余数据，服务器资源消耗严重

资源损耗

- 集群间大量数据复制，消耗大量网络、ETL乃至人力资源

成本高，使用难！
传统MPP数据库已经无法支撑大型银行的大数据建设和发展！！

龙跃MPP DB核心架构



➤ **存算分离统一存储**，实现全集群数据共享，消除数据冗余和数据孤岛，减少数据冗余和服务资源30%以上。

➤ **无状态计算层**，计算和并发能力线性扩展，实现从100节点到20000节点的规模扩展

➤ **云原生架构设计**，实现一键集群创建、启停、升级、扩缩容、故障自愈等功能，运维效率提升10倍以上

龙跃MPP DB——管理控制台

权限管理

- 多租户/用户管理

集群生命周期管理

- 创建、删除、扩缩容、升级、启动、停止

IaaS资源交互和调度

- 基础设施资源调度，包括计算、存储和网络资源等

自动化/智能化运维

- 自动化安装部署
- 监控、告警
- 故障自愈

The screenshot displays the '数据仓库' (Data Warehouse) management console. The left sidebar contains navigation links: 首页 (Home), 数据仓库 (Data Warehouse), 公网IP (Public IP), 联系人组 (Contact Groups), 任务 (Tasks), 弹性网卡 (Elastic Network Cards), 系统管理 (System Management), and 租户管理 (Tenant Management). The main panel shows a table of data warehouses with columns for ID, Name, Type, Tenant, Available Zone, Status, and Remarks. There are buttons for '新建元数据服务' (New Metadata Service), '新建数据仓库' (New Data Warehouse), '启动' (Start), and '停止' (Stop). A search bar and filter options are also present.

ID	名称	类型	租户	可用区	状态	备注	创建时间
7d23b430-c499-4990-a7...		计算服务		建行云-自用1区	运行中	PAB_compute_pro(5...	2021-4
5fb21091-d446-45d4-8c8...		计算服务		建行云-自用1区	运行中	NEPIS_compute_etl(...	2021-4
6facb244-c023-49b9-a66...		计算服务		建行云-自用1区	已删除	NEPIS_compute_etl	2021-4
b94318e3-d1aa-43e6-ae2...		计算服务		建行云-自用1区	已删除	MAA_compute_etl	2021-4
c100929a-2b47-4c20-85c...		计算服务		建行云-自用1区	运行中	PFPC_compute_etl	2021-4
8d29e395-096a-474f-a81...		计算服务		建行云-自用1区	运行中	SIA_compute_etl(128)	2021-4
d4924446-8f2c-4238-be7...		计算服务		建行云-自用1区	运行中	AIS_compute_pro(1...	2021-4

龙跃MPP DB——元数据服务

调度层

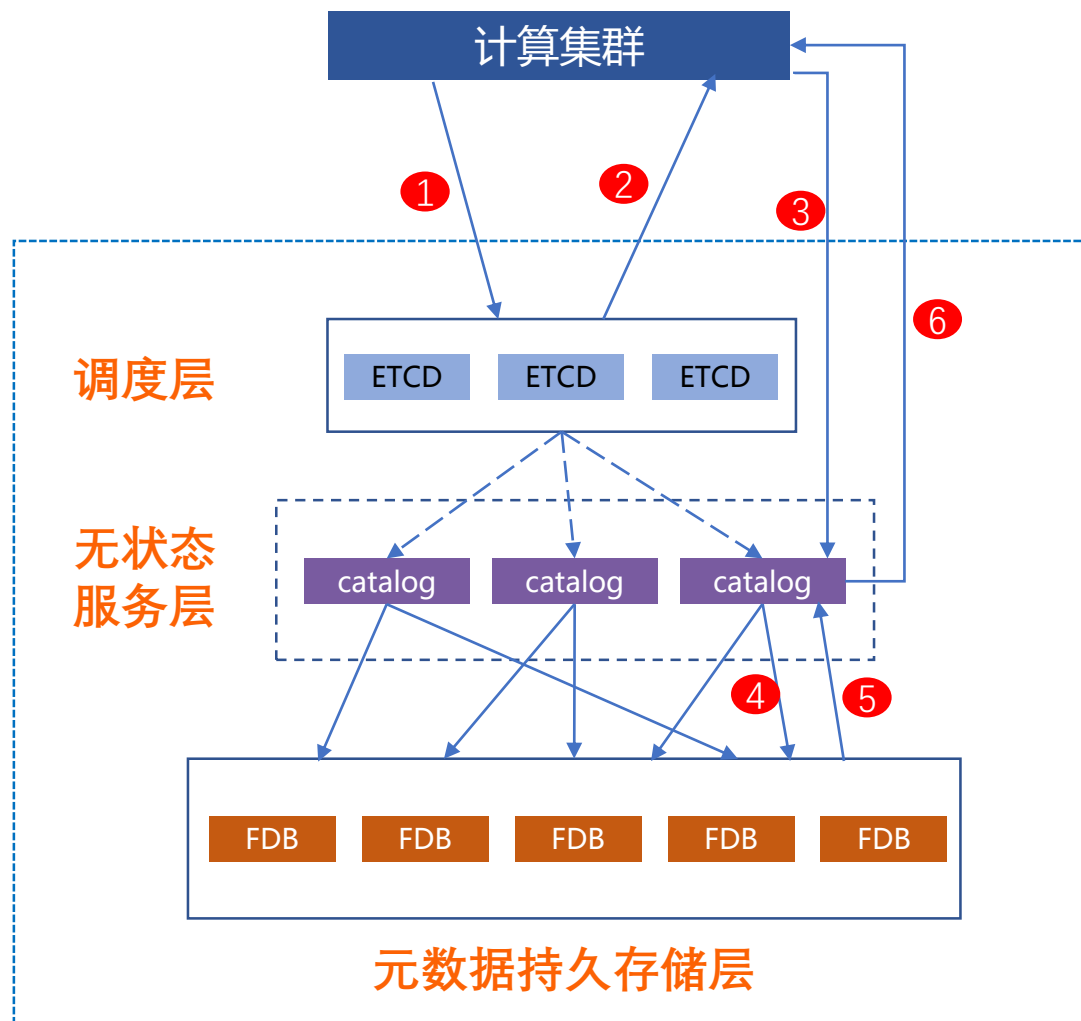
- 服务发现和监控
- 负载均衡

无状态服务层

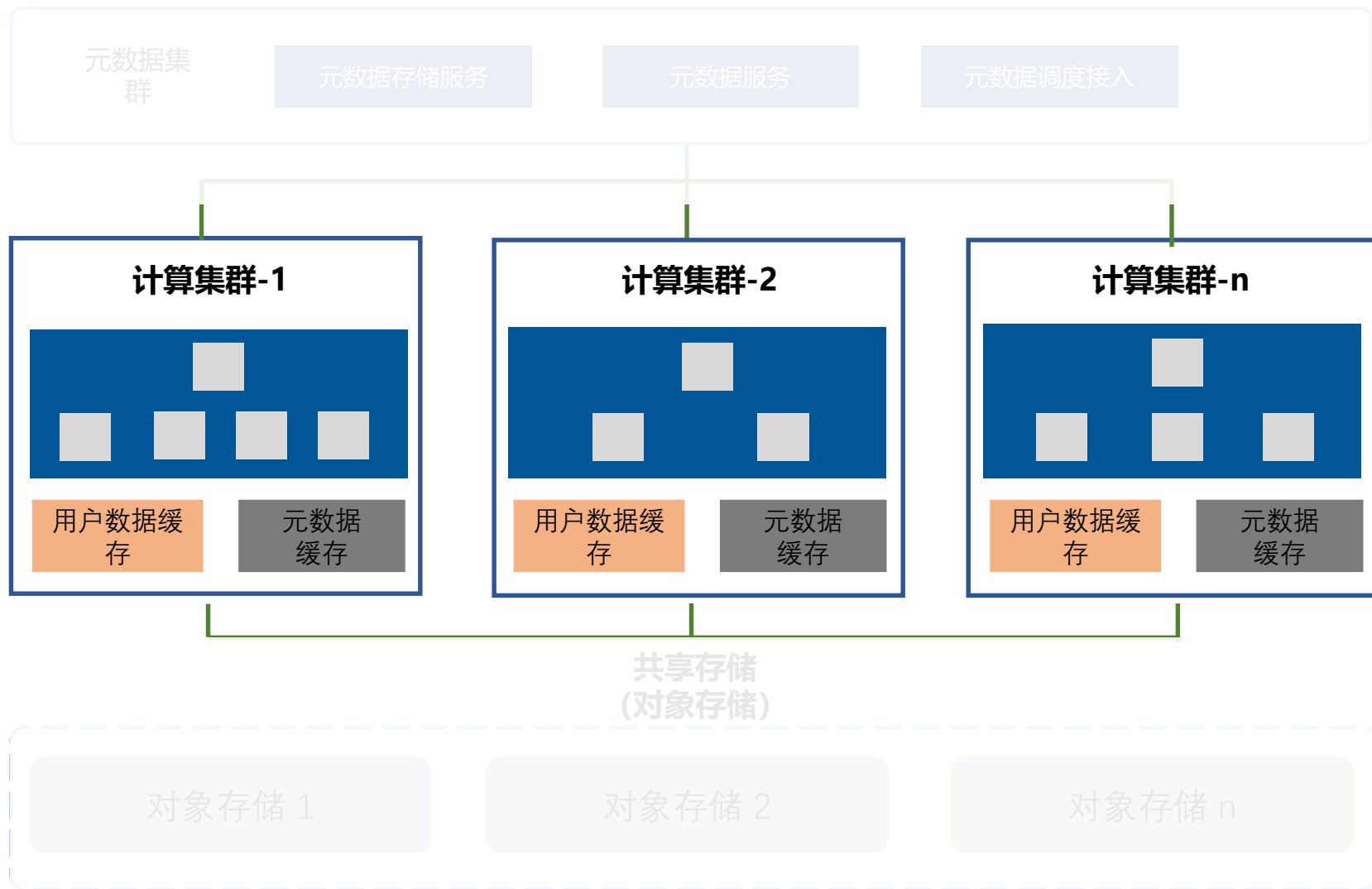
- 服务层由一组服务节点组成，每个服务节点其实是无状态的服务进程，负责接收和处理计算集群的元数据请求；

元数据持久存储层

- 元数据持久化存储层，存储数据字典、统计信息、文件映射等信息



龙跃MPP DB——无状态计算层



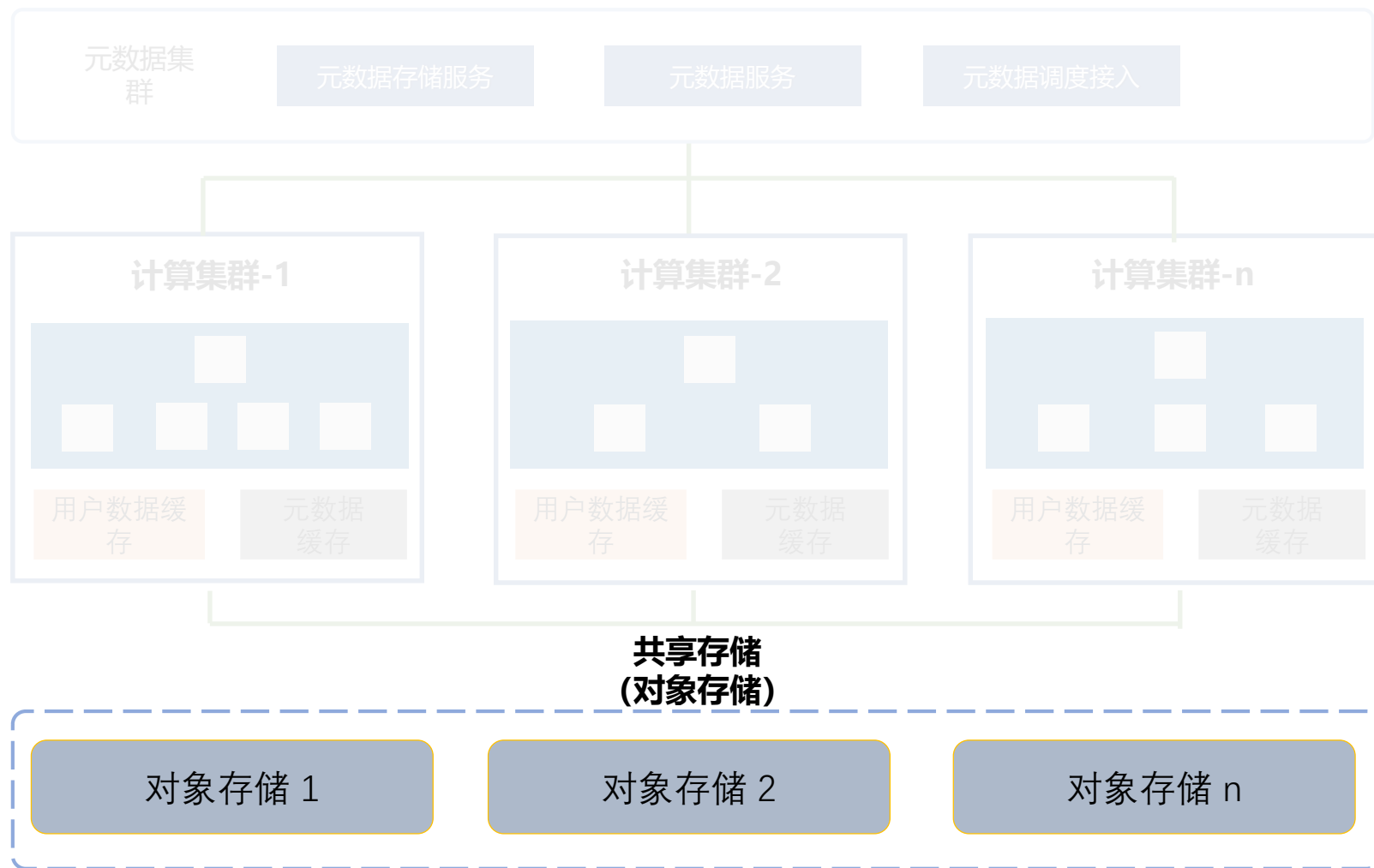
资源灵活分配

- 按需创建、删除、扩缩容
- 集群间资源完全隔离
- 作业可在不同集群建灵活调配
- 并发能力线性扩展

缓存服务

- 本地SSD作为缓存介质
- 元数据缓存
- 用户数据缓存

龙跃MPP DB——共享存储层



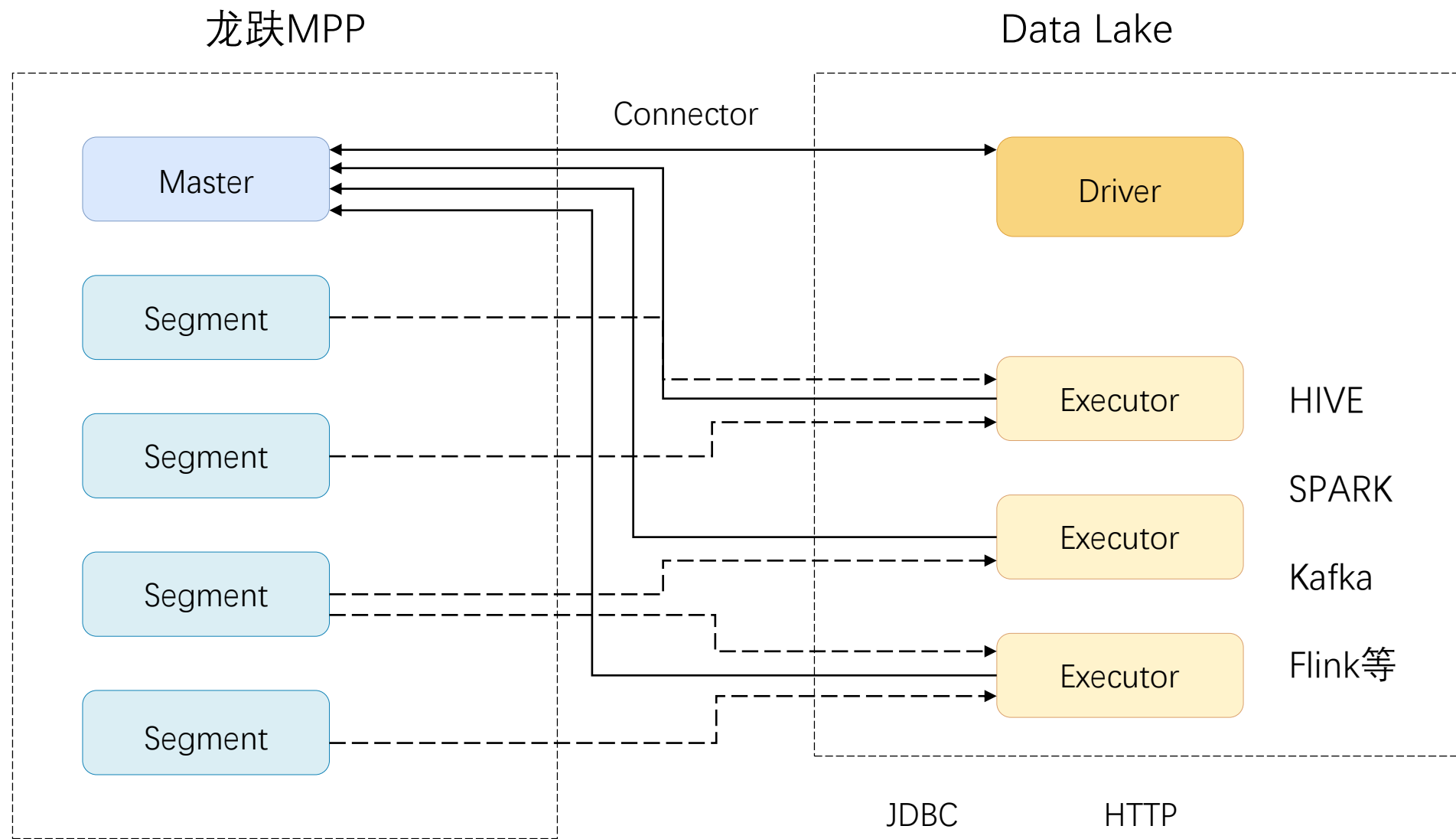
使用对象存储作为数据持久化存储

- 支持100亿文件对象，200PB以上数据
- 使用标准Restful API，支持高并发访问
- 99.99%以上的可用性
- 99.999999999%以上的数据持久性

存储访问优化

- 多桶存储
- 列存+压缩

龙跃MPP DB——丰富的外部计算引擎连接器



龙跃MPP DB——银行级安全可控

数据仓库/数据湖

龙跃MPP DB

服务器

 HUAWEI

 浪潮
inspur

 Great Wall 软件

.....

云平台

 建行云

 金山云

 腾讯云


 中国通信服务
CHINA COMSERVICE


 HUAWEI

 QINGCLOUD 青云

.....

CPU

 Kunpeng

 Phytium 飞腾

 HYGON
中科海光

操作系统

 KINGSOFT
麒麟软件

 统信软件
UNIONTECH

存储

 QINGSTOR®

 XSKY®

 杉岩数据
SANDSTONE

 腾讯云

.....

网络

 HUAWEI

 H3C
数字化解决方案领导者

 Ruijie 锐捷
NETWORKS

 紫光集团
TSINGHUA UNIGROUP

.....

龙跃MPP DB核心技术总结

- 1

架构云原生

融合云计算，架构上实现了存储、计算、元数据分离，摆脱了传统MPP数据库的各种架构限制和制约。
- 2

计算存储分离

共享统一的元数据和数据存储，以增加计算集群的方式，横向扩展集群并发计算能力，突破传统MPP并发限制。
- 3

应用松耦合

数据库资源和应用解耦，根据计算集群的工作负载，和变化的应用需求，灵活、动态调配计算集群。
- 4

分析多样化

在MPP并行计算架构上，集成GIS、Python等组件，支持SQL分析、机器学习、时空分析等多种分析形式。
- 5

秒级扩缩容

数据保存在共享存储上，使用一致性Hash 分布方案，避免新增/减少节点数据迁移。
- 6

自愈高可用

计算集群节点无状态，数据保存在共享存储上，没有Mirror节点，实现分钟级新节点恢复。
- 7

湖仓一体化

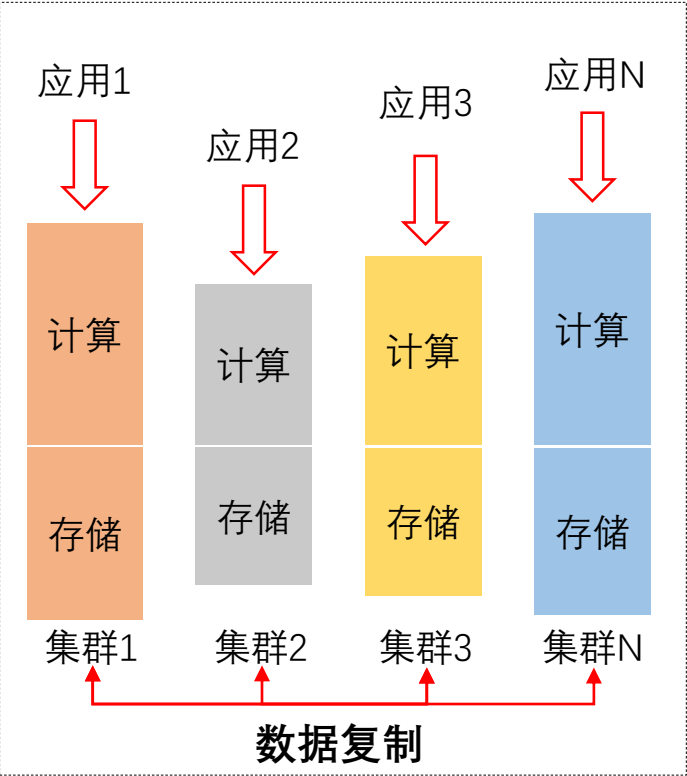
统一数据存储，形成企业级数据视图，统一元数据管理，严格事务一致性，以融合架构支持湖、仓平台建设
- 8

技术自主可控

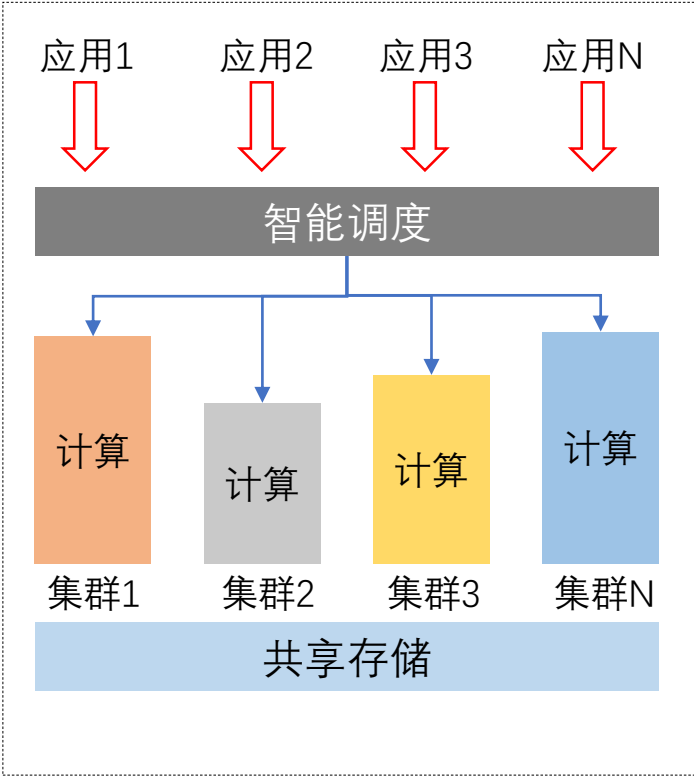
自主开发；支持鲲鹏、海光服务器，麒麟等国产化操作系统，平台软件实现自主可控。



基于龙跃MPP DB的超大规模数据仓库建设解决方案



传统MPP应用解决方案

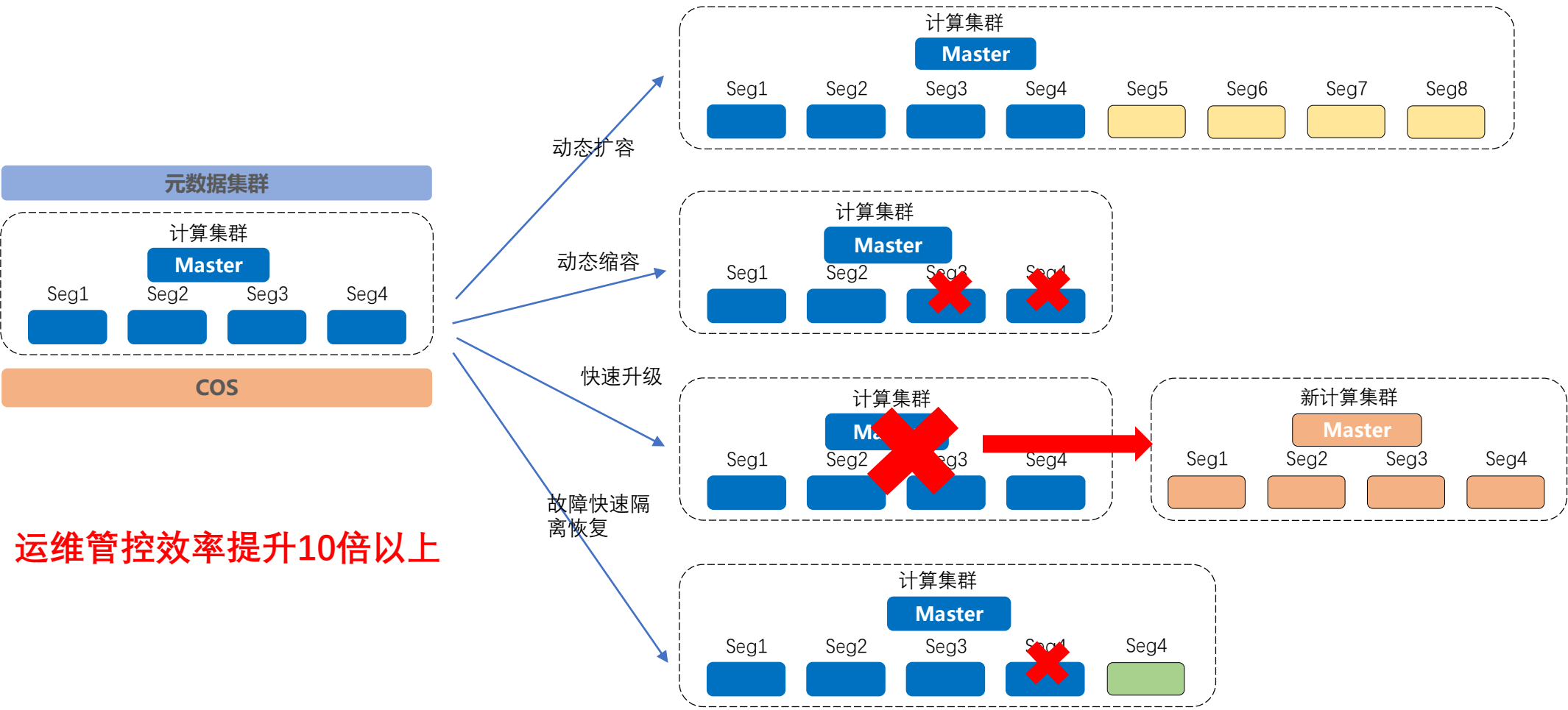


龙跃MPP DB应用解决方案

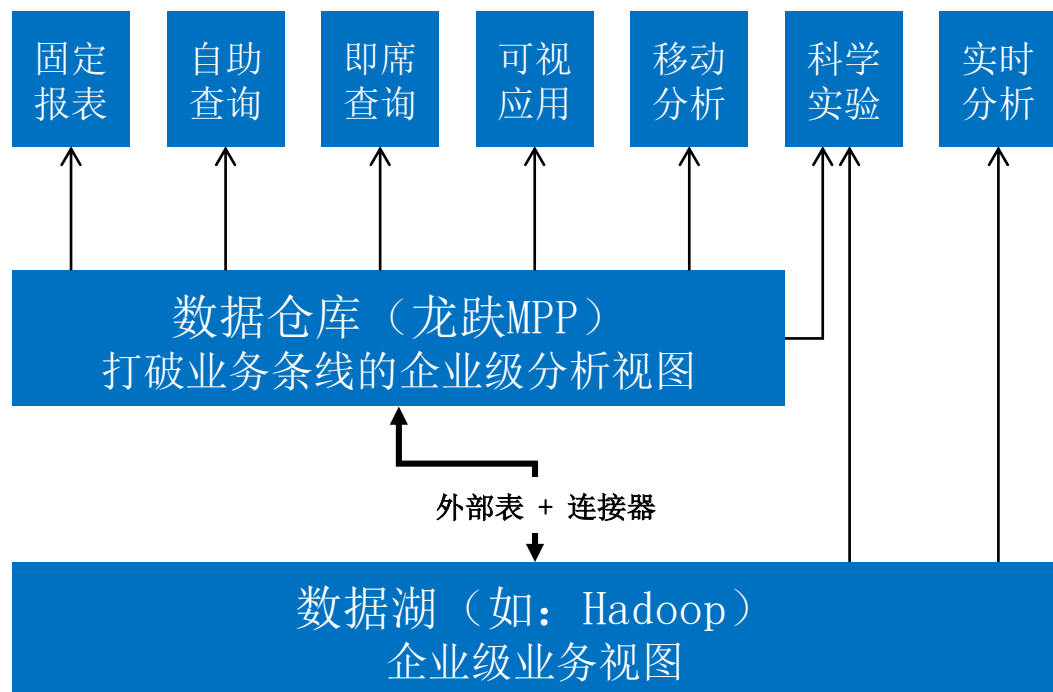
- 完全消除复制作业，减少作业压力20%以上；
- 消除数据冗余，减少数据存储30%，节省硬件资源30%以上；
- 缩短数据使用链路，平均作业完成时间减少1-3小时

	传统MPP	龙跃MPP DB
数据复制	大量集群间数据复制	完全消除数据复制
数据冗余	大量冗余数据	无数据冗余
资源灵活调配	计算和存储资源固定	根据应用需求，实施分配和调整应用资源
作业动态调度	每个集群运行作业基本固定，无法动态调整	作业可以根据负载需求，在不同集群间动态调整
扩展能力	不超过1000个节点	轻松突破10000节点以上

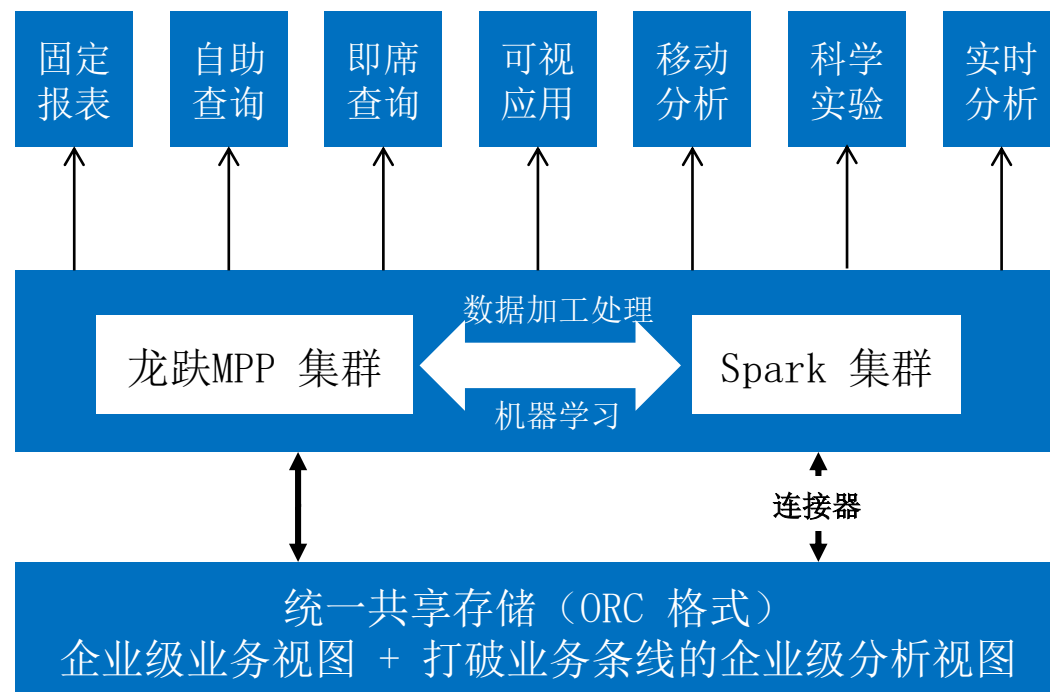
基于龙跃MPP DB的超大规模数据仓库建设解决方案



基于龙跃MPP DB的湖仓一体解决方案



- 数据湖和数据仓库共存的两层混合架构，两者定位不同，各司其职，通过龙跃MPP外部表和连接器集成；
- 数据湖管理各种形态的业务原始数据，一方面为数据仓库提供数据，另一方面支撑数据科学实验和实时查询分析场景；
- 数据仓库打破业务条线，从分析角度整合数据，累积历史，支撑固定报表、灵活查询等传统BI分析，以及部分数据科学实验需求。



- 湖仓一体的融合架构，数据以ORC格式统一存储在共享存储，实现湖仓业务数据和元数据真正融合；
- Spark执行数据计算处理后，将数据按约定的格式写入到共享存储中，同时更新MPP的元数据信息，MPP直接使用写入数据。Spark通过相应的连接器，直接访问MPP写入的数据，不需要重新加载，直接进行机器学习；
- 湖仓一体平台同时支撑所有业务场景。

目录

01 基于龙跃MPP的超大规模数据仓库建设

02 龙跃MPP技术架构和带来的应用革新

03 基于龙跃MPP的大数据未来技术展望

未来技术规划

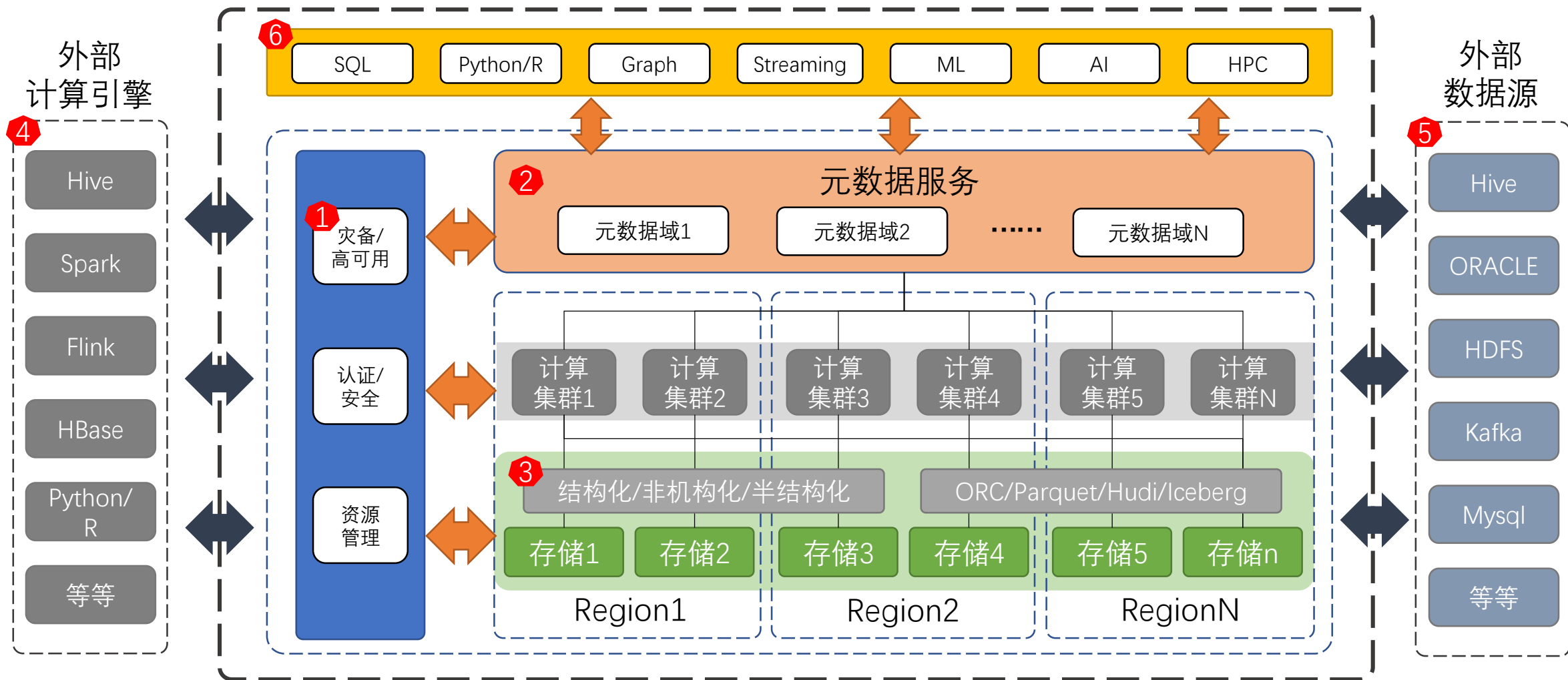
- 基于龙跃MPP DB，我们实现了超大规模数仓建设，资源线性扩展；实现了计算与存储分离，数据不冗余；实现了云原生架构，一键式部署一键式运维。
- 那么展望未来大数据技术发展，我们希望大数据计算与存储技术，能够提供什么样的能力？

安全可靠，服务不中断，数据不丢失

能够高效接入、存储、管理和计算不同类型、不同格式的数据

具备或者能够集成多样的计算能力，满足不同场景的应用需求

多维度资源弹性伸缩，多租户负载管控，全场景故障隔离和自动恢复



1 具备跨区域的高可用部署和灾备恢复能力，数据多区域多版本存储，服务不中断数据不丢失；

2 多租户管理能力，元数据服务具备负载管控、故障隔离的能力，全资源多维度服务弹性伸缩；

3 存储和管理结构化/半结构化/非结构化数据，涵盖主流开放式存储格式和数据湖格式；

4 外部计算引擎能够并行高效的查询和使用龙跃MPP的数据，并将结果存回库内；

5 接入丰富外部数据源，利用龙跃MPP计算能力满足外部数据计算乃至联邦计算的需求；

6 对外能够提供SQL、AI、图、时序、流等计算能力，全面覆盖上层大数据应用需求。



让金融科技尽其所能
thanks