

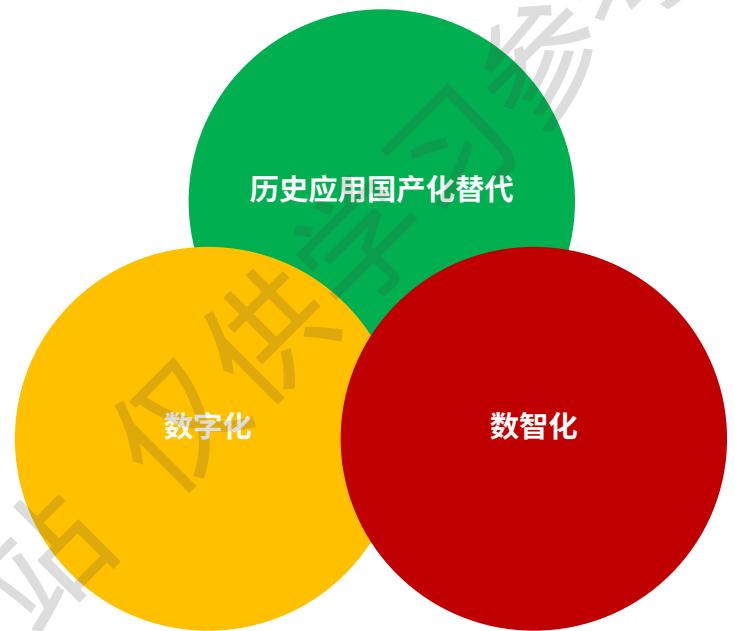
从“专用割裂”到“融合共生”： AI时代的数据库架构演进

分享嘉宾：张昊

应用演进趋势

KING BASE | 金仓社区

对比维度	企业业务演进		
	信息化	->数字化	->数智化
核心内涵	流程电子化 : 将线下业务流程搬到线上, 用IT系统支撑和固化现有作业	业务数字化 : 将物理世界的业务和信息转变为数据, 利用数字技术重构业务流程和商业模式	决策智能化 : 基于数字化产生的海量数据, 利用AI、大数据等技术进行智能分析、预测和决策
核心目标	提升效率 : 通过无纸化、自动化, 减少人工错误, 提高内部管理和流程执行效率	创新与增长 : 创造新的价值点、新的业务模式和新的收入来源, 提升用户体验和市场竞争力	价值最大化 : 实现智能决策、业务自动化和资源最优配置, 降本增效和发现新机遇
关注焦点	流程 (Process) : 关注如何用技术优化和固化内部流程	体验 (Experience) : 关注如何用技术重塑用户/客户体验和商业模式	价值 (Value) : 关注如何从数据中挖掘洞见, 创造新知识, 并驱动行动
数据角色	结果 : 数据是业务流程产生的副产品, 用于事后记录和查询	生产资料 : 数据是重要的生产元素, 是业务创新和优化的基础燃料	核心资产与驱动引擎 : 数据通过AI模型转化为知识和智能, 直接驱动决策和行动
数据价值	记录价值	要素价值	知识价值
技术角色	工具(Tool)	平台(Platform)	驱动引擎(Engine)
思维模式	支撑思维	重构思维	智能思维
典型应用	财务记账软件	滴滴/ Uber (连接车、路、人)	抖音/淘宝的个性化推荐
	仓库库存管理系统	智慧餐厅 (扫码点餐、后厨联动)	工厂的预测性维护
	人事档案数据库	产品全生命周期管理	金融领域的智能风控
技术基础	计算机、软件、 数据库 、网络, 如ERP, CRM, OA, 财务软件等	云计算、 大数据 、物联网、移动互联网、5G技术深度融合, 连接物理与数字世界	机器学习、深度学习、大数据分析、RPA及 人工智能 成为核心驱动力
组织形态	部门级应用 : IT部门主导, 系统通常是烟囱式、孤立的	企业级转型 : 需要业务与技术深度融合, 是“一把手工程”, 往往是跨部门的协作	生态级智能 : 组织整体是数据驱动的, 并能与外部生态进行智能协同



数据库处于 AI 应用技术栈的记忆+简单认知环节，需增强理解、学习、推理、生成能力

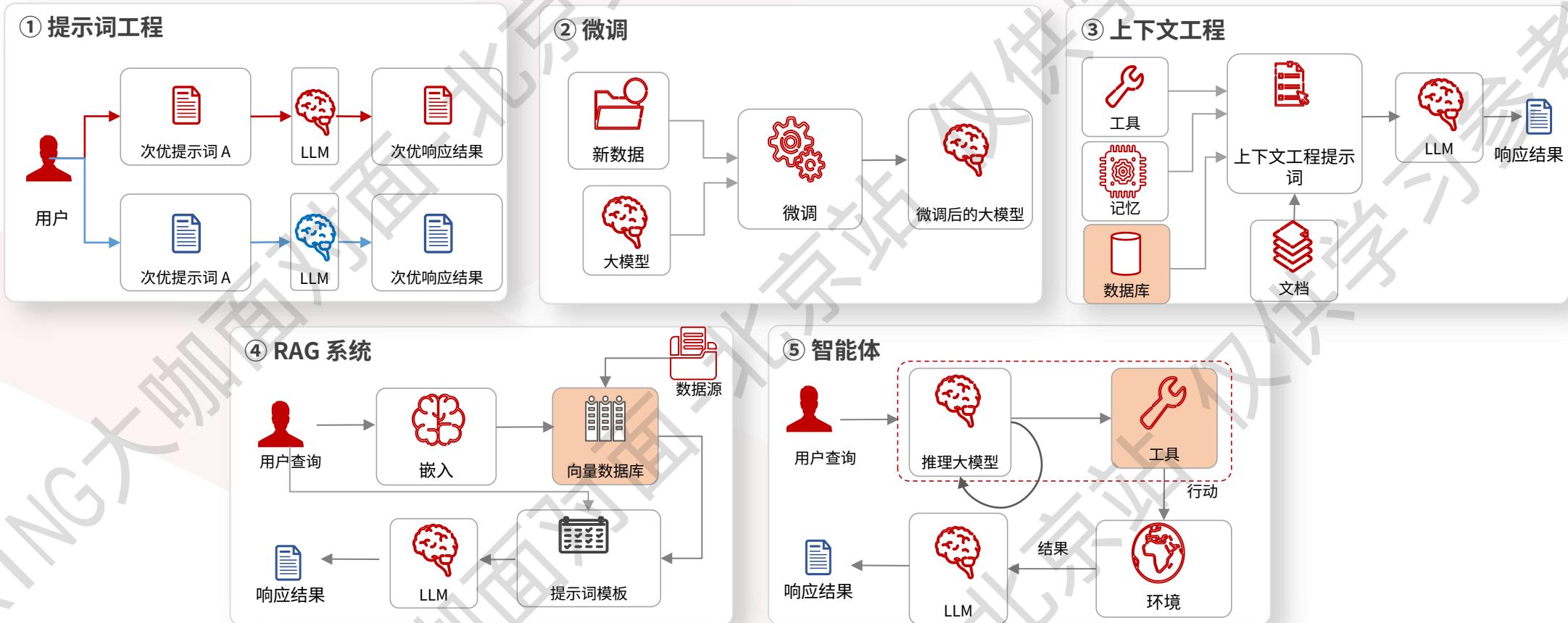
KING BASE | 金仓社区

人体	人脑子系统	主要功能	类比AI	类比数据库
人体	生命维持	内稳态调节	能源、 内外部通信	能源、 内外部通信
		基本生命维持		
	感知	触觉	CV、 语音识别	请求解析
		视觉		
		听觉		
		多种感觉整合		
	记忆	记忆形成	向量数据库	数据存取
		记忆巩固		
		记忆存储		
	语言	语言理解	LLM	数据计算
		语言产生		
		逻辑分析推理		
		计划决策、问题解决		
	高级认知	运动	MCP Server	DB MCP Server
		情绪		
		注意力	暂无	暂无
	奖励	动机		

用途	备注
认知	理解 语义处理：辨别语义中的上下文、细微差别、微妙之处
	统计分析 从样本数据推断总体特征
	多维分析 从多个维度或便利中提取有意义的信息
	分类 将数据划分到预定义的类别中，用于识别和判断。
	聚类 将数据归类为若干类别
	预测 预见未来发生的事件或趋势，用于风险评估、需求规划等。
	学习 学习范围包括：基础语料、上下文
	推理 进行多步推理，更好执行逻辑、数学、编程等任务
表达	生成 创造新的内容，如文本、图像、音乐、视频等。

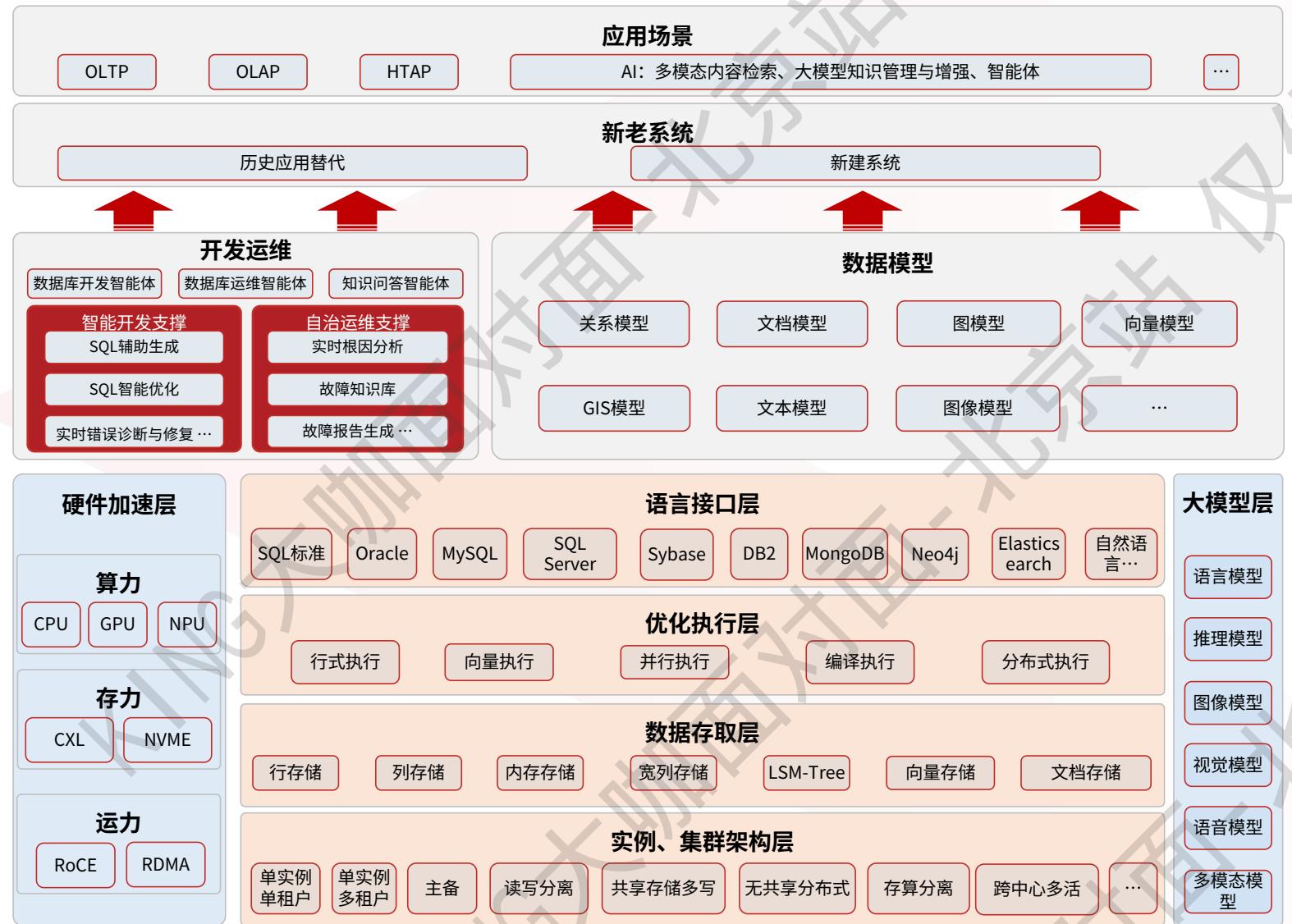


AI 工程技术栈及数据库的位置



通过数据库与大模型协同，为应用提供更强的理解与推理、生成能力

KES 融合数据库架构



面向用户需求的“四个一体化”架构	
一体化层次	用户需求
全历史应用一体化平替	应用程序无缝迁移至国产数据库
多应用场景一体化承载	一个数据库平台支撑覆盖OLTP、OLAP、HTAP、AI等多种应用场景
多模数据一体化处理	收敛技术栈，降低应用复杂度和成本 减少库间数据同步，降低同步开销
开发运维一体化管理	AI赋能开发运维：智能增效，降低系统使用门槛、提高系统效率

基于技术实现的“六大融合”技术体系	
融合体系	技术路径
多语法体系融合	原生语法无损兼容，语义与事务一致性保障，多模型语法统一支持
多执行引擎融合	多执行引擎透明融合，自动获得最优执行路径
多存储引擎融合	数据形态按需存储（行存、列存、无模式等） 跨模型事务协同，加速应用开发与计算效率
多集群架构融合	多集群架构满足不同级别的可用/业务连续性、性能扩展性、成本需求，最大化投资价值
软硬协同融合	软硬协同，系统根据硬件环境动态调整执行策略
AI大模型融合	与大模型协同，为应用提供更强的理解与推理、生成能力

多语法体系融合：助力历史应用一体化平替+多模 数据一体化处理

KING BASE | 金仓社区

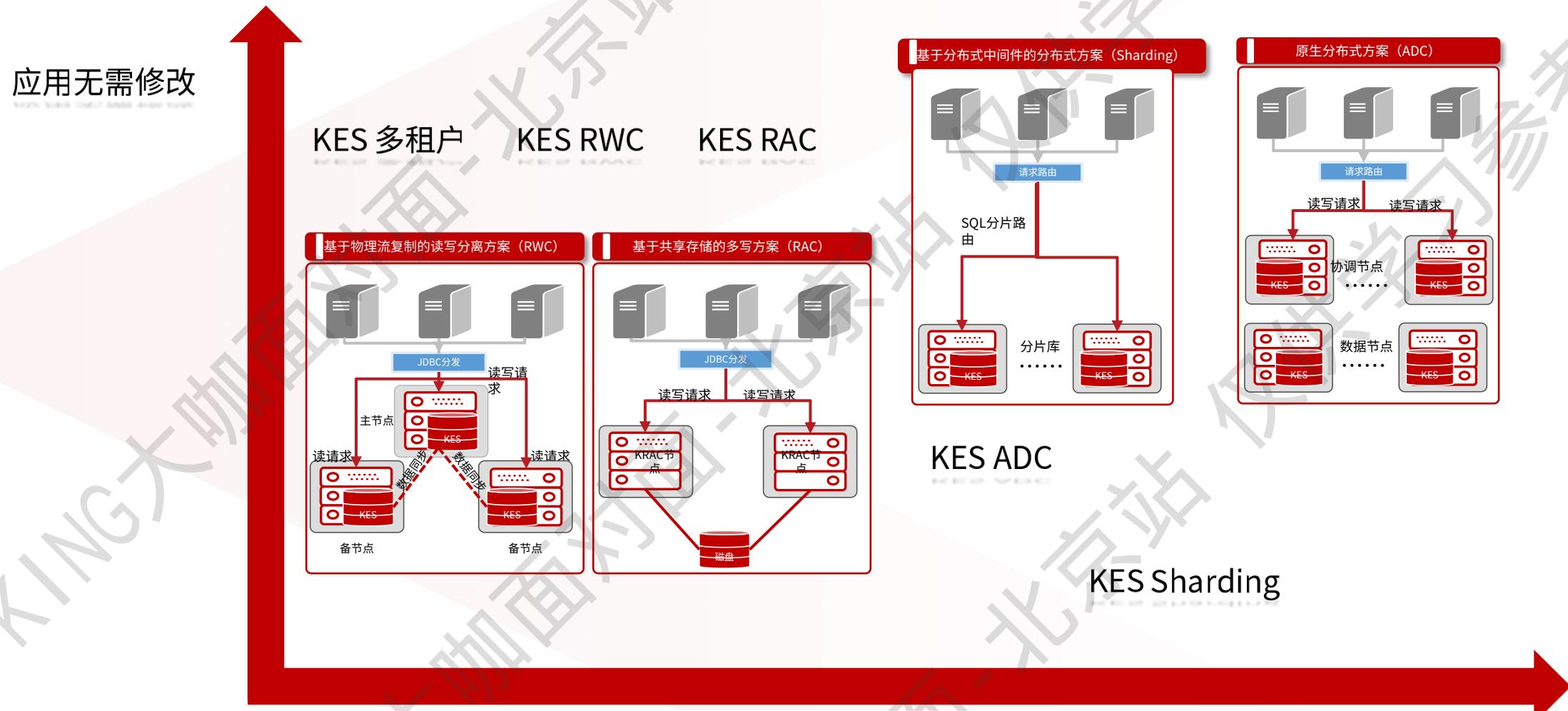
可插拔体系架构

- > 插件式体系架构
- > 多种兼容模式，支持参数选择
- > 支持全新数据库，快速兼容开发

SQL语言	数据类型 伪列 条件表达式 数据库对象 对象操作 数据类型 参数
PLSQL语言	控制语句 执行语句 事务处理 异常处理
SDK	内置函数 内置系统包 内置视图
客户端编程接口	C、Java、Perl、Php…

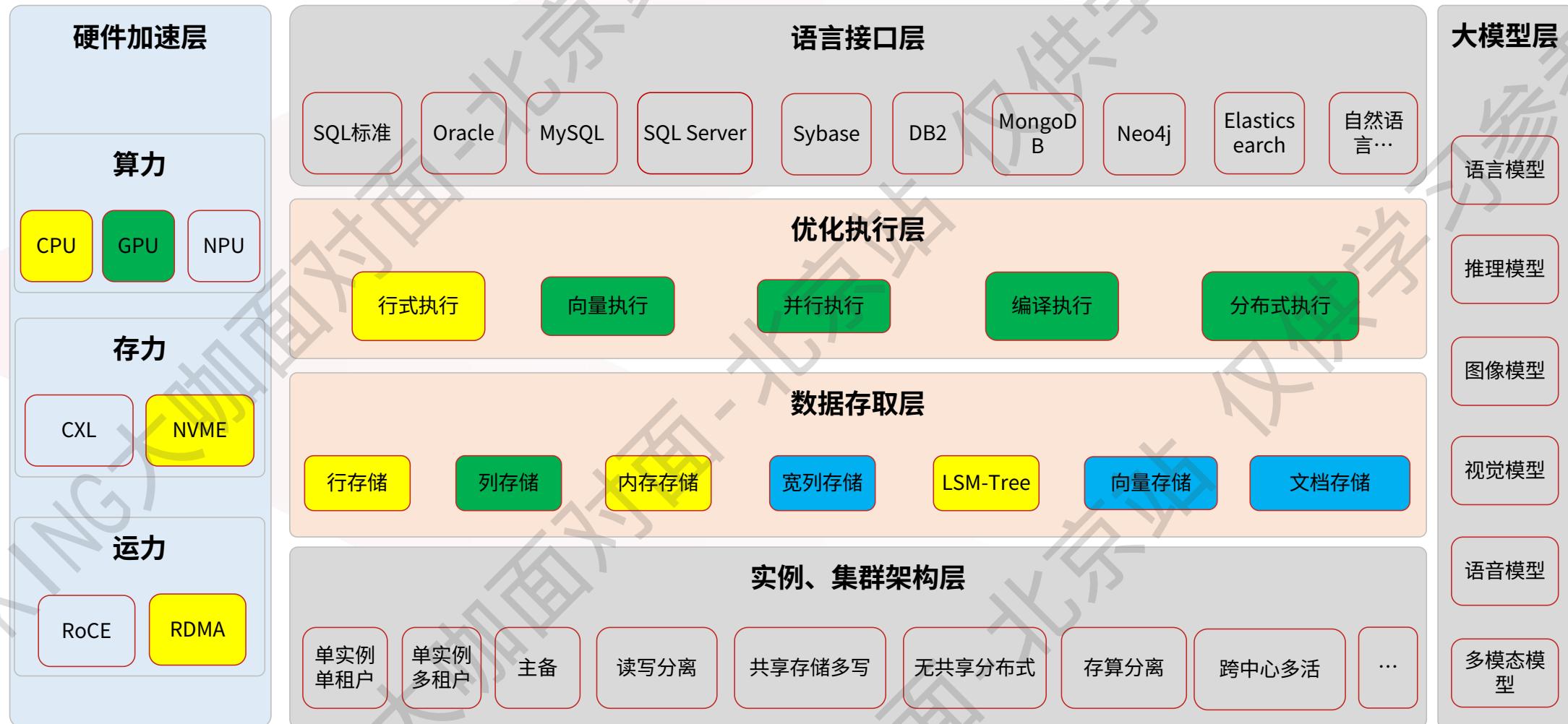


多集群架构融合：助力多应用场景一体化处理



多执行引擎+多存储引擎+软硬协同融合：助力多应用场景一体化处理+多模数据一体化处理

KING BASE | 金仓社区



智能查询优化器

- > **逻辑查询优化**: 内置150+优化规则，自动改写低效SQL，降低执行路径筛选空间
- > **物理查询优化**: 更精确的数据分布统计信息，助力准确估算执行代价，选出最优路径
- > **自适应执行计划**: 基于执行结果反馈自动修正执行计划

匹配全诊断过程的支撑能力

- > **感知**: 对标Oracle，低影响、全覆盖系统负载、运行状态、资源使用等的感知能力，全面掌握系统运行情况
- > **诊断**: 基于DBTIME的问题诊断树，准确定位瓶颈问题
- > **建议**: 基于专家经验、模拟验证的优化建议，助力DBA找出最佳解决路径

应急处置手段

- > **HINT**: 对标Oracle的Hint能力，提供表访问路径、表连接顺序、表连接算法、并行执行算法等多种Hint，指导优化器固定执行计划
- > **QueryMapping**: 不修改应用，将问题源SQL语句映射成等价目标SQL语句，改进系统运行效率
- > ...

开发运维智能体

- > **开发智能体**: SQL辅助生成、SQL智能优化、实时错误诊断与修复，加速SQL开发效率与效果
- > **运维智能体**: 自动巡检、故障知识库、实时根因分析、故障报告生成、修复建议，实现预测性防护、问题精准定位和故障自治恢复

高质量的指标数据



精准的运维知识



丰富的运维经验



强大的推理大模型



答案

DB与AI大模型融合
DB+AI深度融合

案例：多模数据一体化处理

典型场景需求

查询某市过去30天内与雨天+分支道路+追尾形式类似的交通事故频发（至少3次）的5个热点区域

返回结果包含

- 事故数量
- 事故详情
- 最早和最后发生时间
- 与预期特征的相似度

关键数据模型

- 某市热点区域：固定地理范围，以及细化区块/网格 -> GIS数据模型
- 雨天+分支道路+追尾形式：相似度检索 -> 向量数据模型，按文本描述的语义相似度查询
- 区域内各事故详情：文档聚合 -> 文档数据模型

```

WITH city_boundary AS (
  SELECT geom AS city_geom
  FROM areas
  WHERE name = '西安市'
  LIMIT 1
), filtered_accidents AS (
  SELECT
    a.accident_id,
    a.accident_time,
    a.location,
    a.details,
    a.embedding,
    ar.area_id,
    ar.name AS area_name,
    a.embedding <= '[0.12037, 0.49821, 0.42626, .....]'::vector(768) AS vector_distance
  FROM accidents a
  CROSS JOIN city_boundary
  JOIN areas ar ON ST_Contains(ar.geom, a.location)
  WHERE ST_Within(a.location, city_boundary.city_geom)
  AND a.accident_time >= NOW() - INTERVAL '30 days'
  AND a.embedding <= '[0.12037, 0.49821, 0.42626, .....]'::vector(768) < 0.5
)
SELECT
  fa.area_id,
  fa.area_name,
  COUNT(*) AS accident_count,
  json_agg(fa.details) AS aggregated_details,
  MIN(fa.accident_time) AS first_accident,
  MAX(fa.accident_time) AS last_accident,
  AVG(fa.vector_distance) AS avg_vector_distance
FROM filtered_accidents fa
GROUP BY fa.area_id, fa.area_name
HAVING COUNT(*) > 3
ORDER BY accident_count DESC
LIMIT 5;

```

说明：

- 事故的描述文字通过嵌入模型生成向量（本例中为 768 维），保存在表 a 的 embedding 字段中；
- 查询输入的文本“雨天在分支道路上追尾”，也通过上述模型生成嵌入向量（内容过长，以 [0.12037, 0.49821, 0.42626,] 示意）；

GIS计算

明确事故所在区域

向量搜索

事故特征与查询特征的余弦距离

文档处理

分区域聚合事故详情清单

分组排序

分组计算热点区域排序

GIS：国家生态保护红线监管平台项目

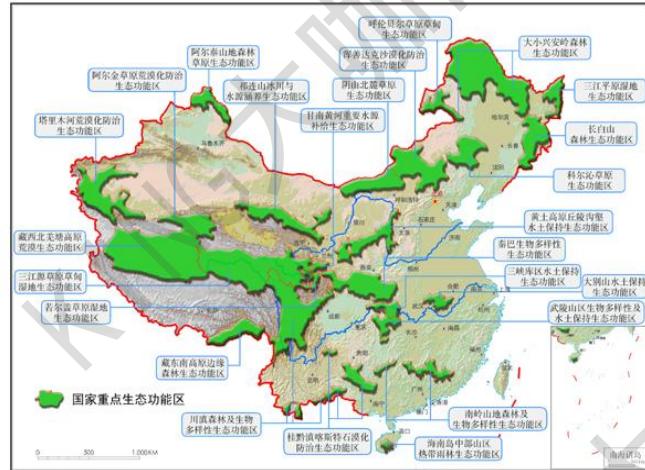
KING BASE | 金仓社区

■ 项目背景

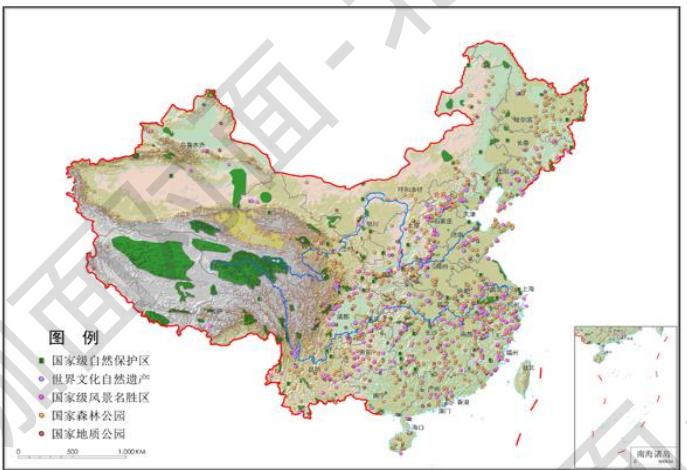
- > 自然资源部负责生态红线的划定和勘探勘界；生态环境部负责生态红线的监管和督查。通过卫星遥感、大数据、人工智能等技术手段，监控生态红线的破坏和减少情况。
 - > 项目应用环境：ArcGIS，X86，Linux。

支撑情况

- > 数据迁移：采用特有的空间数据全量离线迁移工具，实现Oracle Spatial到人大金仓空间数据库的快速自动化平滑迁移。
 - > 生态适配：全面适配ArcGIS平台，保障本项目中基于ArcGIS的众多GIS应用无缝迁移并高效运行。
 - > 稳定性：数据库并发连接1000+，在大并发高吞吐量业务压力下，数据库稳定运行。
 - > 性能：提供经优化的空间数据索引，并对空间函数进行深度优化，使得空间数据处理性能提升80%。



国家生态功能区分



国家禁止开发区域展示

48小时 完成 **3TB+ Oracle Spatial 空间数据迁移**

空间数据迁移

空间
处理
性能
提升
80%

**1000+
并发连接
7*24
稳定运行**

MongoDB：福建某地市电子证照共享服务系统

项目介绍

2024

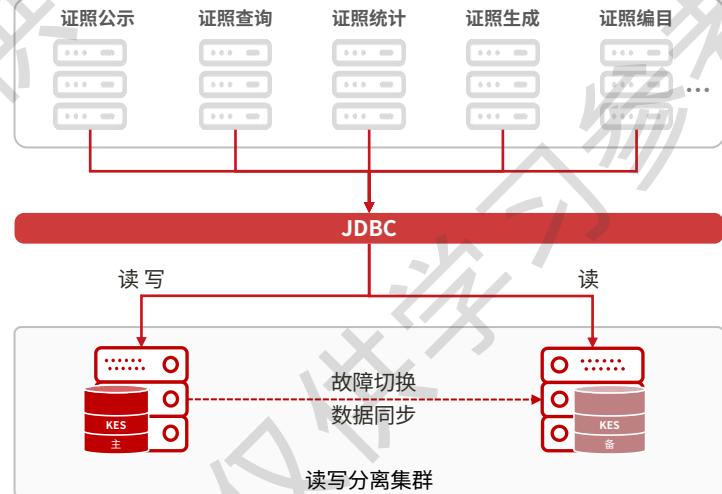
- > 基本情况：通过电子证照共享服务提升项目，实现三明电子证照高效供给，持续深入证照标准化水平、数据质量等方面拓展，提高政务服务水平，构建新型智慧政务体系。
- > 业务情况：依托三明市云计算中心政务云平台、政务信息网及政务外网等基础设施，在现有电子证照的架构基础上，通过升级电子证照应用服务系统、电子证照管理系统、新建证照链系统，原系统国产化改造等建设内容，提升服务渠道、内容、体验。
- > 架构情况：原系统采用x86芯片+Linux系统+MongoDB 2节点主从集群；新系统采用海光芯片+FusionOS操作系统+2节点 KES 读写分离集群。

用户痛点

- > 痛点1：MongoDB 存储文档数据，而国产关系型数据库需严格遵循表结构规范等诸多方面差异，整体改造成本高。
- > 痛点2：系统服务覆盖 500 余家单位，高峰期并发量达 1000+ 连接数，一些高频操作的业务场景响应延迟偏大。
- > 痛点3：核心数据需在周末时间指定迁移窗口内完成数据迁移、校验等工作，且需确保数据零丢失。

解决方案

- > 平滑替换：金仓多模数据库 MongoDB原生协议兼容，实现零代码平替。
- > 性能优化：金仓读写分离集群突破高并发瓶颈，结合场景化优化，提升系统承载能力。
- > 数据迁移：定制迁移工具助力高效迁移、数据自动比对校验。



价值呈现

原生协议兼容
MongoDB数据库

复杂查询场景
优化后性能提升30%

高效平滑迁移
异构数据迁移

案例：历史应用一体化平替

案例名称	对象数量	plsql 典型的行数（复杂对象）	典型嵌套层次	用户代码是否修改	迁移和验证时间
融先科技	约1450个	1.50个左右的包(Package)，代码量大的包达到3万行代码，一般在1万行左右 2.700个左右存储过程 (Procedure) ，大部分500-1200行代码之间	10-15层	是	1个月
中汇亿达	约4200个	1.258个包，包最大5万，一般在5000-10000行； 2.函数，存储过程代码量一般在100-1000行之间 3.总共约200万行代码	一般5-6层 复杂的超过10层	否	2周
HKJC	约7100个	1.786个包，包代码量一般在1000-1500行； 2.5184个触发器，代码量一般在60-150行之间 3.450个存储过程，平均每个存储过程在230行代码	一般5-8层	否	半周
西南油气田分公司	35119个	1.其中包对象246个，一般有2000-5000行代码。 2.存储过程，函数992个，一般有300-1000行代码 3.触发器33347个，一般有30-100行代码	一般5-6层	否	1周
上海国际港务	2081个	1.208个包代码，一般有5000-10000行代码 2.1185个存储过程，单个代码量主要集中在100-300行之间	一般5-6层	否	2周
高法	1452个	最大的8647行，平均500+以上	一般5-6层	否	3周
上期CTP交易系统	140个	1.全部为包对象 2.大部分包代码量在1万-2万行代码； 3.最大的包(Package)达到8万行代码； 4.有包的函数(Function)个数超过2000个。	至少6层	否	2周

案例：多应用场景一体化处理 (TP)

应用系统	系统用途及核心访问特征	新系统架构	新硬件配置		并发规模		数据规模 (TB)	事务吞吐量 (TPS)	跑批性能	响应时间	RPO	RT0
			组件	配置	用户数	数据库连接数						
嘉实基金 TA 系统替换	基金业核心系统，包含账户、交易、基金产品清结算等业务，如基金登记过户、认购、申购、赎回、转托管、基金转换等	4套1主1备读写分离集群	CPU	Kunpeng-920 128C	亿级	2000~8000	70+	20000+	3h完成42类跑批任务涉及7TB数据	普通业务秒级核心交易毫秒级	0	秒级
			内存	1TB								
电网营销	抄表、核算、账单、收款、客户服务管理等	3套1主1备读写分离集群	CPU	Kunpeng-920 192C	5000万	10000+	100+	10000+	2h完成5000W	毫秒级	0	秒级
			内存	1TB								
北京市政一卡通清算	1.5亿张卡，10分钟入库一次	1主1备读写分离集群 + 1 节点容灾	CPU	X86 2路36C 72T	2000万	2000+	12+	5900+ (早高峰3h)	<2h完成126万笔任务	毫秒级	0	秒级
			内存	512GB								
邮政集团 OA	全国大集中架构，支撑邮政、邮储全集团员工的日常办公、档案等业务	2套1主3备读写分离集群 (历史数据单独存储)	CPU	鲲鹏920 128C	30万+	3000~10000	10+	55600+	N/A	毫秒级	0	秒级
			内存	1T*2 +512G*2								
上海高法核心业务	23+法院，206 个核心业务模块	4套一主两备集群	CPU	鲲鹏920 128C	3W+	9500+	14+	10000+	<8h完成5000+案子审批报结	<2s	0	秒级
			内存	512GB								

案例：多应用场景一体化处理 (AP)

应用系统	系统用途及核心访问特征	新系统架构	新硬件配置		并发规模		数据规模 (TB)	事务吞吐量 (TPS)	跑批性能	响应时间	RPO	RTO
			组件	配置	用户数	数据库连接数						
北京移动数据共享平台	内部生产数据共享、大数据量处理	40节点	CPU	X86_64 32C	10	150	100	150	准实时 每小时完成 100+跑批	<1h内完成固定 跑批任务	0	分钟级
			内存	256GB								
北京市健康宝平台	疫情数据统计准实时展示	10节点	CPU	X86_64 32C	8	100	50	100	准实时 每15分钟完成 60+跑批	<15分钟内完成 固定跑批任务	0	分钟级
			内存	256GB								
LJ大数据领导管理领域基础平台	大数据量处理和安全性要求高	10节点	CPU	Kunpeng-920 虚拟化16C	5	50	30	50	<6h完成 500+跑批	报表业务秒级， 核心跑批小时级	0	分钟级
			内存	64GB								
中国科学院国家天文台信息基础平台	天文观测数据碰撞分析、大数据量处理	23节点(2*10+3)	CPU	X86_64 16C	10	180	60	180	每6小时完成 300+跑批	报表业务秒级， 核心跑批小时级	0	分钟级
			内存	256GB								
中华人民共和国海事局QTDXC	进出口贸易货物统计	11节点(北京5节点、上海6节点)	CPU	X86_64 64C	15	200	30	200	<8h完成 800+跑批	报表业务秒级， 核心跑批小时级	0	分钟级
			内存	256GB								
东莞市智慧大脑	所有政府部门数据统一汇总展示	10节点	CPU	X86_64 64C	50+	400	40	400	<8h完成 1000+跑批	报表业务秒级， 核心跑批小时级	0	分钟级
			内存	256GB								
国家药监总局“互联网+政务服务”平台	药品监管统计分析	8节点	CPU	X86_64 16C	10	100	20	100	<6h完成 300+跑批	报表业务秒级， 核心跑批小时级	0	分钟级
			内存	128Gb								

案例：开发运维一体化管理（湘财证券多核心系统）

项目介绍

2024

- > 基本情况：随着《关于构建更加完善的要素市场化配置体制机制的意见》《证券期货业科技发展“十四五”规划》等政策出台，数据安全被提升至国家战略高度。证券行业作为金融核心领域，需防范数据泄露风险，实现数据存储与处理的自主可控。
- > 业务情况：湘财证券国产化项目涵盖客户服务与营销、交易与投资、合规与监管、运营管理和运维与管控等多个业务领域，其中重要核心业务系统包括：镒链场外衍生品交易管理系统、恒生资管 TA 系统、IPO 网下发行系统、恒生投保报送系统、卡方算法总线系统、信用交易中心和营销中心等超过50套应用系统。
- > 架构情况：新系统采用KES单实例、主备集群架构模式，通过金仓KEMCC企业级统一管控平台进行运维管理。

用户痛点

- > 痛点1：用户担心数据库厂商在面对大规模数据库管理场景时，无法提供统一的数据库运维管控方案。
- > 痛点2：用户担心在大规模应用场景下，数据库一旦发生故障，很难快速全面的定位故障原因。
- > 痛点3：用户担心应用系统使用多套不同版本数据库，数据库补丁升级工作复杂。

解决方案



价值呈现

统一运维管理

100+KES
单机、集群

集中监控

可视化集中监控
故障快速定位、分析

补丁升级

单点、批量下发
不同版本下发

下一代 AI 数据库：认知更深入、行动力更强、使用更简单

KING BASE | 金仓社区

面向领域	角色	数据库发挥的作用	核心功能
DB→Data DB4AI	数据库 数据	结构化数据、 非结构化数据 语义数据分析	GIS、文档、全文检索数据库
	数据库作为 MCP Server	作为大模型的 手脚	基于自然语言的查询处理（自集成 NL2SQL/PLSQL 模块或借助大模型）
	数据库作为记忆体	存储、利用大模型的训练结果，作为大模型的 记忆体	KES MCP Server 向量数据库
	数据库增强生成	利用存储的知识库、知识图谱，作为大模型的 外脑	图数据库
	数据库深化认知	对数据库内的数据进行 高级分析	内置数据挖掘、机器学习算法，实现 回归, 规划, 预测, 推理

以应用驱动产品发展，打造DB4AI应用最佳方案实践

编号	关键能力	阶段成果
1	面向应用开发人员、DBA、数据科学家等数据库的使用者（ 面向我司外部用户 ），打造外部智能体	3个外部用户智能体+知识图谱
2	将经验泛化到销售管理、数据库内核开发工程师、测试工程师、QA、产品经理等岗位，应届生等亟需快速学习提升的人群（ 面向我司内部岗位 ），打造内部智能体	6个内部岗位智能体+知识图谱
3	利用智能体构建经验，沉淀智能体构建方法论、方案与最佳实践	1个智能体构建方法论



THANKS

成为世界卓越的数据库产品与服务提供商

