

金仓迁移全生命生命周期最佳实践

01-迁移方案背景

在当前国际/国内背景下，国产化替代工作已经不是“要不要做”，而是已经转变为“如何做”的问题了。

金仓原厂人员每年直接为客户部署近万套数据库，直接服务客户上线近 2000 个系统，在多年的信创及替“0”迁移实践中积累了大量的项目实践经验。

本课件即以金仓客户现场迁移实操经验为积淀，提炼、萃取形成金仓迁移最佳实践。行业专业 DBA 可以通过此课程迅速建立：从已使用源数据库向金仓数据库迁移的工作能力——在工作实践中完成自己单位的数据库国产化替代工作。

一、迁移难点

当前，数据库迁移工作属于“引入期”阶段，很多从业人员因对此项工作了解有限，面对如下问题，普遍存在“不会用、不敢用、不愿用”情况。

1、海量数据的平滑迁移。

原系统数据库存量数据经过长时间的使用，通常已经到 TB 级，每日增量数据达数百 GB，技术团队必须在极短割接窗口期内完成海量数据的不停机平滑迁移，业务不中断/少中断。

2、高负载迁移后场景支撑。

原系统数据库通常已经在高负载下运行，很多项目最大日交易量可达数千万笔，TPS 几千笔/秒，迁移后性能需要满足应用要求。

3、对源数据库语法支持。

原系统数据库（如 Oracle、MySQL、SQL Server 等）的语法兼容问题——各数据库厂家在自己数据库发展历程中，会逐步形成自己“特色”体系结构、语法（SQL、存储过程、函数、触发器、视图等）、特性等；迁移后目标数据库是否能不修改或少修改数据库相关语法，正常运行。

4、应用适配。

原业务系统采用 B/S 架构或 C/S 架构，在新系统改造升级中，要不修改或少修改应用代码，使客户在无感知情况下完成升级工作。

5、业务停机。

迁移一般会导致原业务系统（短暂）停机，这会对于以 7*24 小时运行的核心系统迁移工作造成巨大的挑战。

二、金仓“三低一平”技术支撑方案

1、金仓在业务实践中，了解到客户普遍存在如下“业务痛点”：

- （1）担心学习成本高、难度大——源数据库和金仓数据库在开发、运维方法是否类似。
- （2）担心应用改造成本高、难度大——源数据库向金仓数据库迁移的兼容性问题：存储过程、函数、触发器、视图、应用内 SQL 等代码是否需要修改。
- （3）担心数据迁移复杂，工作量大，劳心劳力——源数据库累计海量数据和增量数据迁移问题：迁移过程复杂，迁移后数据一致性比对工作量大。
- （4）担心测试无法全面覆盖生产环境——传统测试方法/用例无法完整覆盖生产环境业务功能及负载压力，导致即便测试环境没问题，割接上线后就有 bug，出现重大故障。
- （5）担心停机时间过长，影响业务连续性——重要业务系统停机时间窗口短，无法在规定时间内完成迁移而导致整体迁移任务失败。
- （6）担心国产数据库丢数据、宕机风险——重要业务系统完成迁移割接后，一旦发生故障，无法及时恢复。

2、金仓解决方案：

金仓从如下几个维度来解决客户的担忧。

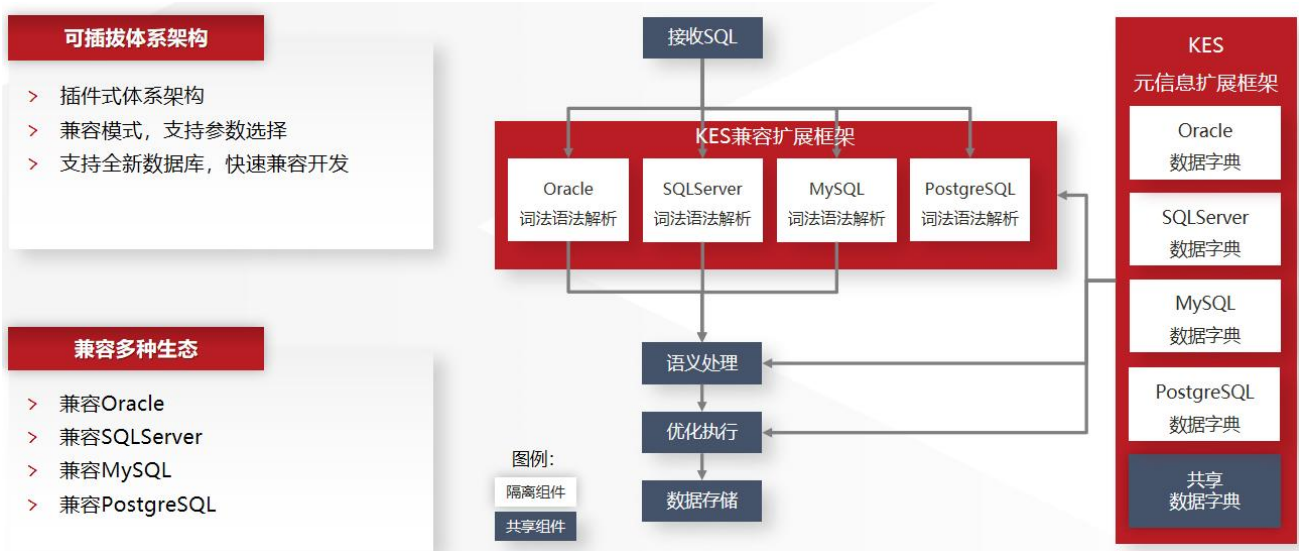
应用适配	应用厂商 无须 修改业务代码，即可运行在国产数据库之上。
数据库开发	开发人员和DBA 无须 从头学习国产数据库的开发和维护方法。
数据库迁移	业务代码和业务数据的迁移简单方便， 支持 不停机迁移。
应用验证	使用负载抓取工具抓取生产环境真实处理过程，并在国产数据库上做全量回归，最大可能提高回归测试覆盖率。
系统上线	切换到国产数据库后万一出现故障，系统可以立即回退到原有系统中，确保 “万无一失” ！

3、金仓“三低一平”技术方案说明：

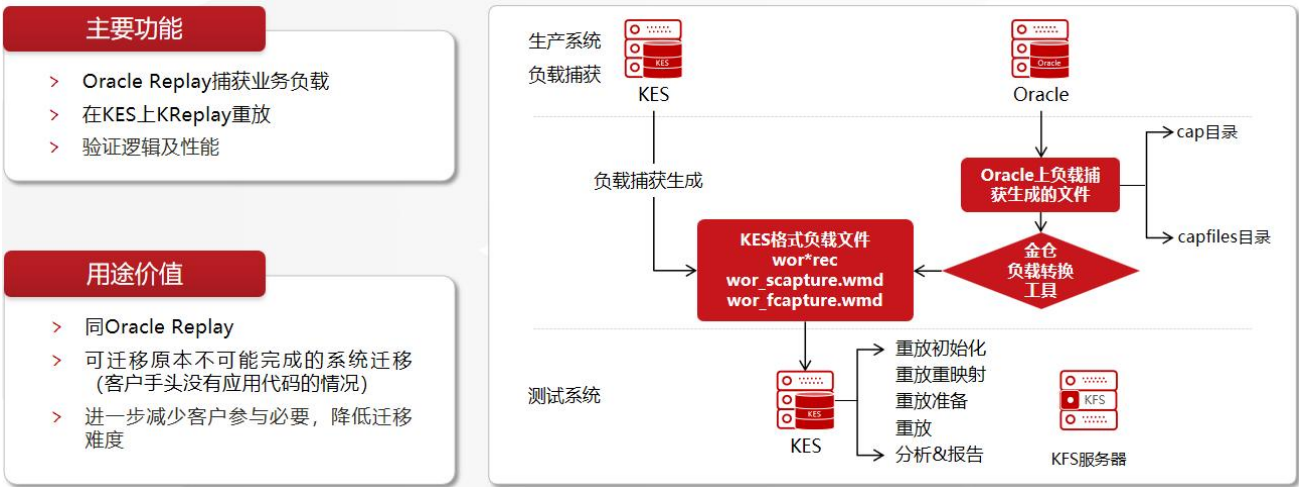
“三低一平”即：低难度、低风险、低成本、平滑迁移，金仓方案可从容帮助客户实现数据库迁移、替代工作。

(1) 低难度：

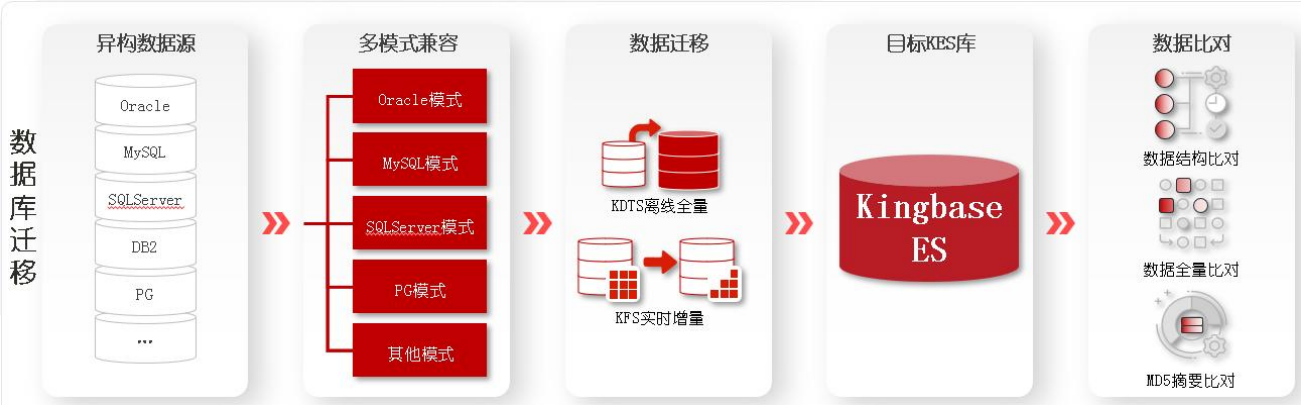
- 多语法原生兼容一体化框架，基本不需要修改 SQL、PL/SQL。



- 基于生产环境的回归验证（Oracle 为例）。

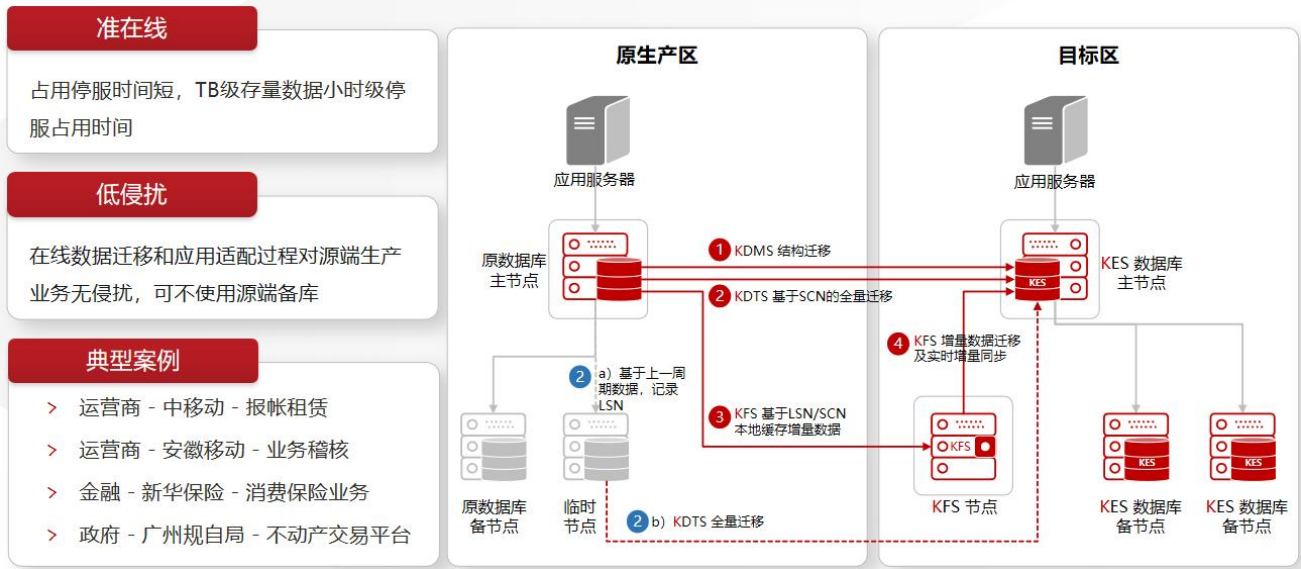


- 全流程工具化，极大降低人工投入。

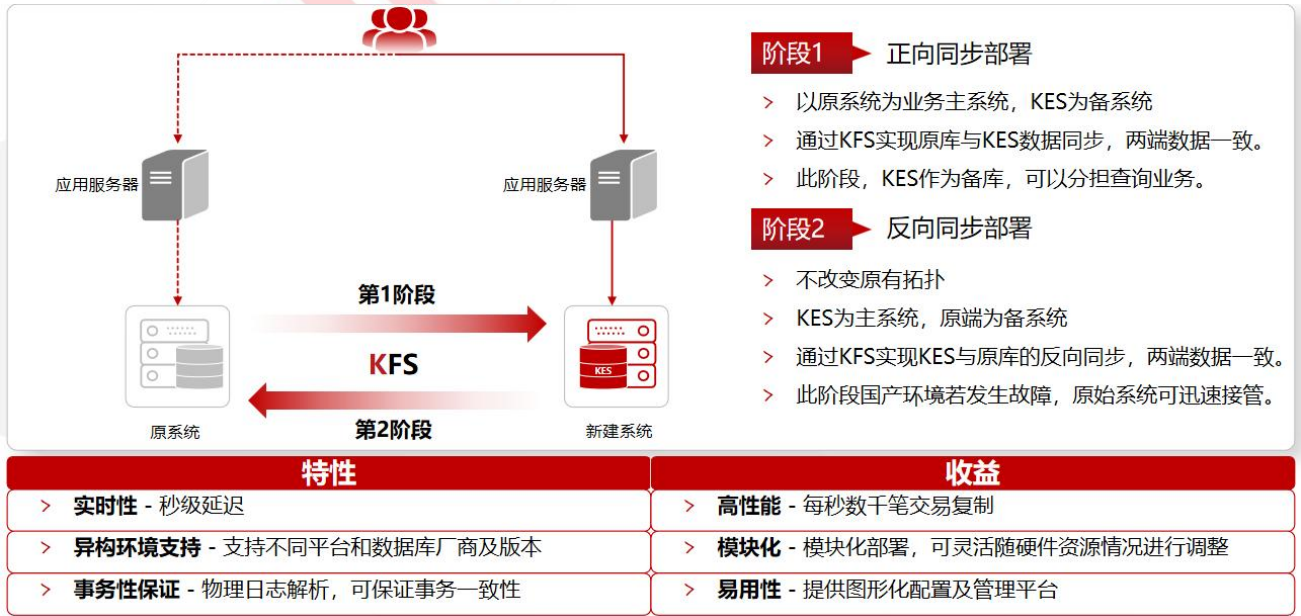


(2) 低风险

- 柔性迁移方案，保障业务连续性。



- 双轨并行方案，保障业务可随时回退到源数据库系统。



- (3) 低成本：
- 迁移成本对比。

应用软件开发商迁移成本评估（单应用）									
数据库厂商	其他数据库 <small>生态不健全，兼容性差，性能一般，需投入大量人员优化</small>			金仓数据库 <small>基本不修改代码：只需更换驱动程序，生态健全，兼容性好，处理性能优异，成熟的产品优化方案</small>			节省工作量（%）		
系统类型	办公应用	一般应用	核心应用	办公应用	一般应用	核心应用	办公应用	一般应用	核心应用
应用改造（人天） 改造+适配+测试	10	20	40	1	1	1	↓62%	↓63%	↓68%
工程师要求	初级	中级	高级	初级	初级	中级			
成本小计（万元）	/	/	/	/	/	/			
集成测试（人天） 功能+性能+稳定性	5	15	30	5	15	30			
工程师要求	初级	初级	初级	初级	初级	初级			
成本小计（万元）	/	/	/	/	/	/			
迁移工时合计（人天）	15	35	70	6	16	31			
割接上线成功率	90%	80%	60%	100%	100%	100%			
保底方案	无	无	无	柔性迁移	双轨并行	双轨并行			
优劣势对比	停机时间长，回切方案不完善，回切存在丢数据风险	停机时间长，回切方案不完善，回切存在丢数据风险	停机时间长，回切方案不完善，回切存在丢数据风险	停机时间短（几乎0停机）	停机时间短（几乎0停机），支持回切，无风险	停机时间短、（几乎0停机），支持回切，无风险			

编者注：

办公应用：客户自用的小型系统，例如：考勤、人力、安全、机房管理等。

一般应用：客户自用或对外提供服务的系统，例如：民政、财务、资产管理等。

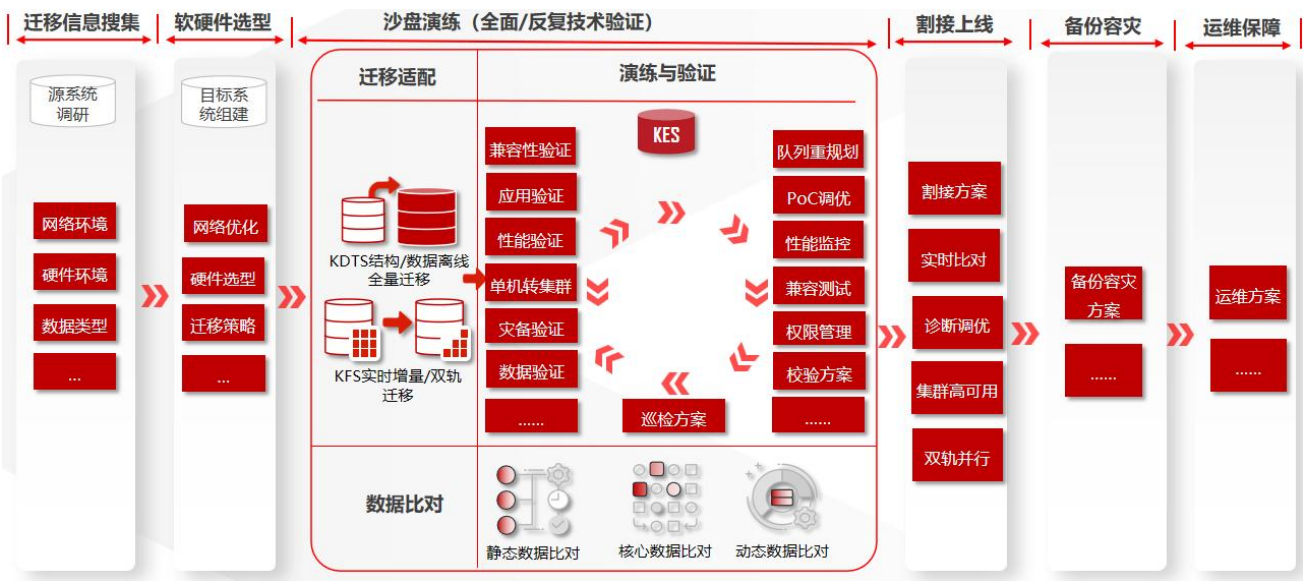
核心应用：主要是对外提供服务的系统，涉及 TOC、TOB 类重型业务，例如：政务办公服务平台、公积金、社保、海关缉私、电子客票、电子口岸等核心应用。

02-金仓迁移全生命周期

一、迁移全生命周期

金仓迁移项目全生命周期如下图所示。金仓迁移生命周期基于金仓公司一系列生态工具构建，立足简单、有效、保质。

在生产中，客户可以基于自己迁移项目的特殊情况，对金仓方法进行适度裁剪，之后再应用于自己的迁移项目管理和落地。



二、迁移信息搜集

任何迁移项目在立项时都需要进行项目迁移信息搜集工作，以便圈定迁移项目范围，并对项目的完成时间、所需花费的成本、完成效果（质量）进行初步规划。

三、软硬件选型

客户在准备做迁移工作前，通常第一考虑优先级是：迁移后的新数据库服务器是否能够满足业务的性能要求。金仓基于多年的实践经验，在本阶段会为客户提供一个简单有效的比对选择的方法。

四、沙盘演练

“沙盘演练”阶段是关系着迁移是否能够成功的关键阶段，客户必须在这个阶段对迁移项目涉及的所有环节逐项进行反复验证、确认，确保万无一失后方可进入割接上线阶段。

五、割接上线

在沙盘演练阶段解决所有迁移技术问题后，才能够进入本阶段。客户在割接上线前需事无巨细的全面考虑可能出现的意外及制定应对措施。如果客户服务器资源充足，可以做一次完整割接上线演练，来确保上线工作万无一失。

六、备份容灾

备份容灾是企业数据管理和业务连续性规划中至关重要的部分，防止数据丢失、维护数据完整性、应对灾难事件、减少业务中断时间,满足合规管理要求，确保业务连续性。

七、运维保障

有效的可落地的运维方案确保数据安全与合规，能够增强系统安全性、提升系统性能，有效保障系统长期稳定运行，在系统出现问题时，也可以高效的进行故障处理与恢复。

03-迁移信息收集

特别说明：本课程使用的所有模板，用户均可用于自己迁移项目，达到对迁移工作有效管理目的。

一、源数据库信息搜集

迁移项目开始前，客户首先需要对源数据库的支撑应用、账户数、日交易量、日数据库事务量、数据库连接数、当前数据量等关键信息进行搜集，这样可以了解迁移项目总体概况，从而做出有效的迁移计划，做到对迁移工作心中有数。

二、迁移数据库的登记台账

在此阶段，金仓在项目中使用如下模板对待迁“源数据库”做台账的登记工作，通过此全面了解迁移项目涉及到的源数据库/应用的全貌。

台账登记汇总信息

数 据 库 IP	应用 IP	应用类型	业务模块	联系人/电话	遗留需求与问题	风险	端口
3.3.3.3	3.3.3.13	XWeb 应用	收款	A 先生/138*****	缺失数据库 备份策略	数据达 3TB	应用：42140 数据库：54321
3.3.3.3	3.3.3.14	Y Web 应用		...	无	无	应用：65530

							数据库：54321
3.3.3.4	3.3.3.15	跑批业务		...	无	无	
x.x.x.x	x.x.x.x	...		D 先生/...	审计方案	无	数据库：54321

编者注：如上模板填写信息为示例，客户应在待迁项目中填写项目的真实信息。

三、登记待迁数据库服务器软硬件

客户基于上一节“台账汇总登记信息”表格信息，逐条使用如下模板搜集、填写每个待迁“源数据库”服务器的软硬件信息。客户在理清源数据库后，可以按如下“金仓经验”建议，对“迁移目标 KES 数据库”的软硬件配置的进行初步选型。

金仓经验：考虑到客户业务的未来可持续的发展，金仓一般建议在“源数据库”的配置和性能基础上，增加至少 20%（是否调成更高？）以上的配置冗余，这样可以在较长一段时间内有效支撑客户业务发展要求。

	源数据库（3.3.3.3）	迁移目标 KES 数据库（3.3.3.6）
系统用途及核心访问特征	高并发写入、复杂 SQL、跑批清算	高并发写入、复杂 SQL、跑批清算
CPU	IBM power8 2*8c 64 线程	intel 2*18c 72 线程
内存	512 GB	768 GB
磁盘	SSD 20TB	SSD（主）SAS（备） 各 25TB
OS	AIX 7.1	麒麟
技术方案	Oracle 2 节点 rac + adg + 2 节点 rac（容灾中心）	KES 三节点读写分离集群（主中心 2 节点，容灾中心 1 节点）
.....		

04-软硬件选型

客户在做迁移工作选型时，通常考虑的最高优先级是：选定的软硬件是否能满足支撑迁移后业务继续有效的运行。

为了简化客户选型工作，金仓在项目里面通常使用如下《金仓数据库服务器配置选型表》进行选型，该表是是金仓公司多年积累的现场经验汇集而成的表格。

此外，IT 行业有多种性能验证评测方案，但相关评测方法通常都较为繁复，实操中技术和管理难度都较大，需要专业人员支持才能够完成相关专项测试。

如果没有机器的情况下如何准备评估和选型对应的硬件规格，这种情况可以寻求硬件厂商的帮助。

设置格式[cheng chen]: 字体: 加粗

删除[cheng chen]: 如果

一、金仓数据库服务器配置选型表

客户在使用此表格选型时，只需要根据下表中的：用户数、日活用户数、平均并发用户数、峰值并发用户数、读写事务比例的区间数据，来选择自己的迁移目标数据库服务器的配置就行。

金仓数据库服务器配置选型表

序号	用户数	日活用户数	平均并发用户数	峰值并发用户数	读写事务比例	数据库服务器（1主N备）				
						8核+16G	16核+32G	32核+64G	64核+128G	128核+256G
1	10-500	350	175	215	172/43	2节点(1主1备)				
2	501-1000	700	350	406	325/81	4节点(1主3备)	3节点(1主2备)			2节点(1主1备)
3	1001-2000	1400	700	779	623/156	5节点(1主4备)		4节点(1主3备)		
4	2001-5000	3500	1750	1875	1500/375	6节点(1主5备)	5节点(1主4备)			4节点(1主3备)
5	5001-10000	7000	3500	3677	2942/735	6节点(1主5备)		5节点(1主4备)		
6	10001-20000	14000	7000	7251	5801/1450	7节点(1主6备)	6节点(1主5备)			5节点(1主4备)
7	20001-50000	35000	17500	17897	14317/3579	7节点(1主6备)		6节点(1主5备)		

编者注：

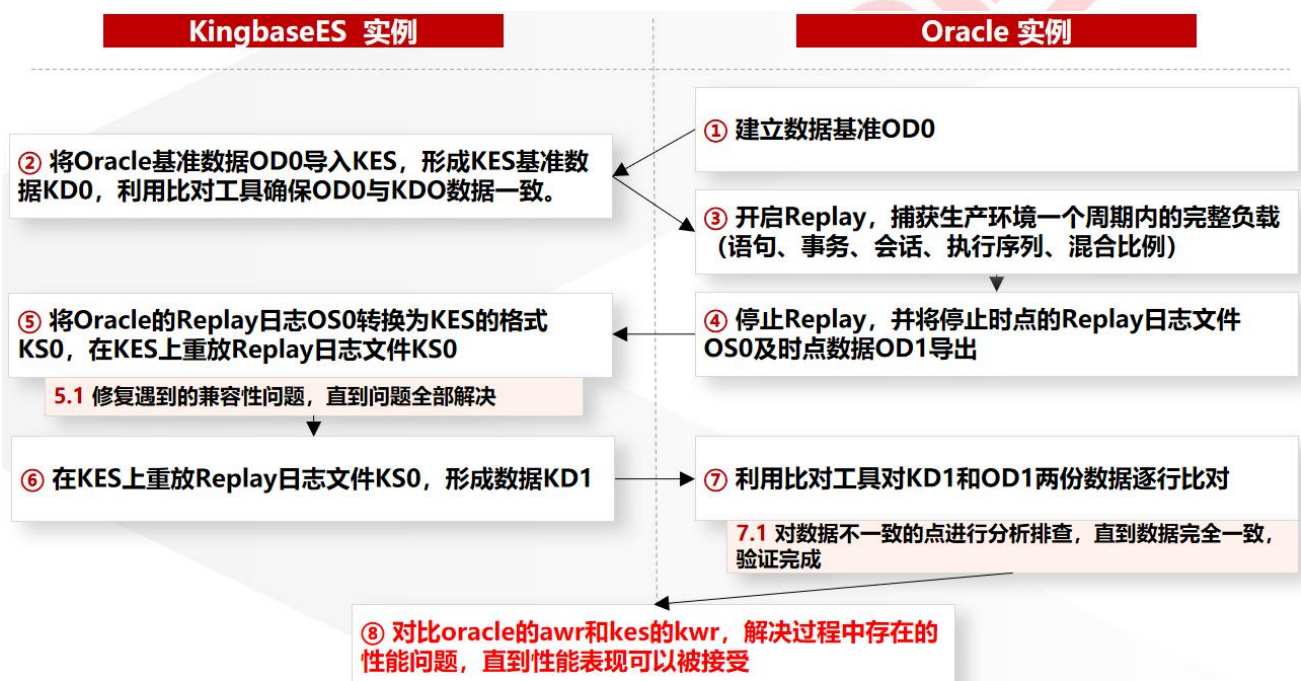
1、日活用户 = 为用户数的 70%。

- 2、平均并发用户数 = N*L/T (N:日活用户数;L:登录的平均长度(用户从登录到退出的时长，此处默认值为 4h);T:指考察的时间段长度(工作日时长，默认 8h))。
- 3、峰值并发用户数 = 平均并发用户数+3*SQRT(平均并发用户数)。
- 4、系统读写比例 = 8:2。

二、通过 KReplay 做兼容性验证

对于业务关键性的无源码的复杂迁移项目（如核心系统），建议需要做在其上运行的应用兼容性测试。

金仓利用其研发的 KReplay 工具进行验证测试，测试环节如下图说明：



三、通过 AWR/KWR 性能验证

对于业务关键型的迁移项目（如核心系统），并且客户有充足的服务器资源，建议客户要进一步利用专业工具进行服务器性能分析、验证、选型。

1、源数据库进行负载采样。

本节以 Oracle 源数据库为例说明，用户可采用 Oracle AWR 做负载采样报告（一般需采集高峰/低谷区段），采集信息涉及：负载分析、实例效率百分比、TOP10 SQL、主机 CPU、主机 IO、主机内存、主机网络、I/O 分析、内存分析等信息。

（1）低谷负载采集。

建议采样时间是凌晨 2:30 到 4:30 期间，采集后用下表记录。一般会按 100%、300%的压力进行推算高倍业务负载支撑下，系统资源可能情况：

凌晨 2:30 到 4:30 期间（业务场景）				
序号	时段	30%（当前业务）	100%业务压力	300%业务压力
CPU 占比	凌晨 2:30 到 4:30 期间	1.6	4.8	14.4
内存占用		32G	96G	288G
事务数		1,267,756	3,803,268	11,409,804

编者注：该时段，100%业务场景下实现 380 万（单一事务数 3 万/分钟）、300%业务场景下 1140 万（单一事务数 9.5 万/分钟）事务数支撑能力。

（2）高峰负载采集。

一般建议时间是客户业务繁忙时段，如下样例表格是以上午 6:30 到 9:30 期间高峰区段进行选择，采集后用下表记录。一般按 100%、300%的压力进行推算高倍业务负载支撑下，系统资源可能情况：

上午 6:30 到 9:30 期间（业务场景）				
序号	时段	30%（当前业务）	100%业务压力	300%业务压力
CPU 占比	凌晨 6:30 到 9:30 期间	1.67	5.01	15.03
内存占用		31.09G	93.27G	279.81G
事务数		59,821,723	179,465,169	538,395,507

编者注：该时段，100%业务场景下实现 1.79 亿（单一事务数 99.7 万/分钟）、300%业务场景下 5.3 亿（单一事务 299 万/分钟）事务数支撑能力。

2、目标金仓库负载采样。

用户可使用金仓公司 KWR 做负载采样报告。在此环节的测试，通常需要客户配合测试，如：提供 1 :1 的测试环境，提供脱敏测试数据，并配合使用 Jmeter、Loadrunner 等工具模拟负载，进行负载测试比对。

（1）低谷负载采集，采样后通过换算关系和 Oracle 比对。

凌晨 2:30 到 4:30 期间（业务场景）				
序号	时段	30%（当前业务）	100%业务压力	300%业务压力
CPU 占比	凌晨 2:30 到 4:30			

内存占用	期间			
事务数				

（2）高峰负载采集。采样后通过换算关系和 Oracle 比对。

上午 6:30 到 9:30 期间（业务场景）				
序号	时段	30%（当前业务）	100%业务压力	300%业务压力
CPU 占比	凌晨 6:30 到 9:30			
内存占用	期间			
事务数				

05-沙盘演练

沙盘演练目的是：遍历迁移过程的所有技术难点，提前解决割接上线的所有可能问题。

一、KDTS 结构/数据离线全量迁移

1、主要适用场景：

- （1）客户待迁移数据库允许停机，停机的时间窗口满足迁移的时间要求。
- （2）原数据库的表对象、非表对象的语法简单。

2、主要流程（流程图化）：

- （1）在源生产数据库做好全量数据备份。
- （2）在中间库恢复备份出来的数据，形成迁移测试环境。
- （3）用 KDTS 工具迁移对象把中间库的数据迁移到 KES 数据库中。
- （4）用 KES 数据库反复进行功能、性能、高可用、灾备等验证，解决凸显问题完成环境验证。
- （5）中间库和 KES 数据比对无误，完成数据迁移模拟工作
- （6）把验证成功方案、环境、数据、脚本等备份出来，做好割接上线前准备工作。

二、KDTS+KFS 实时增量/双轨迁移

1、主要适用场景：

- (1) 客户待迁移数据库不允许停机。
- (2) 关键业务系统，需系统回退保障。

2、主要流程（流程图化）：

- (1) 在源生产数据库做好全量数据备份。
- (2) 在中间库恢复备份出来的数据，形成迁移测试环境。
- (3) 用 KDTS 工具迁移对象把中间库的数据迁移到 KES 数据库中。
- (4) 用 KES 数据库反复进行功能、性能、高可用、灾备等验证，解决凸显问题完成环境验证。
- (5) 使用 KFS 工具进行双轨并行配置，双向切换验证，数据比对无误，完成数据迁移模拟工作。
- (6) 把验证成功方案、环境、数据、脚本等备份出来，做好割接上线前准备工作。

06-割接上线

割接上线阶段是基于沙盘演练的成果来形成割接方案，客户应充分考虑各环节可能潜在的问题，做好系统回退保障，确保迁移万无一失。

一般来说可以采用如下方式缩短迁移时间：

- 1、KDTS 运行机器的内存、CPU 和迁移网段的网络速度非常重要，它决定了数据迁移速度（在沙盘演练环境试迁取得参考数据）。
- 2、对于不怎么变化的表和数据，如字典表等，可以提前迁移。
- 3、对于数据量大的数据表，可以提前或推后迁移数据，一般建议拆分多个 KDTS 任务迁移。
- 4、非表结构对象如存储过程、客户写的函数、触发器等需在沙盘演练的适配阶段验证完成。
- 5、对于大表，索引创建非常耗时，这个因素要充分考虑。
- 6、一般来说，约束最后启用。

一、KDTS 结构/数据离线全量迁移

1、割接主要流程：

- (1) 在既定割接上线时间停止应用服务，使数据不在变化。
- (2) 把验证成功备份出来脚本在迁移目标 KES 数据库执行，完成结构迁移工作。
- (3) 使用 KDTS 工具把源生产数据库的数据全量迁移到目标 KES 数据库，完成数据迁移。

- (4) 进行数据比对，验证无误。
- (5) 配置集群（如割接方案里有集群配置）。
- (6) 应用服务指向 KES 数据库并对外提供服务。
- (7) 进行应用试运行/测试/调优。
- (8) 如一切正常，迁移成功。
- (9) 如发生问题，应用切换到源数据库，继续支持应用服务。

二、KDTS+KFS 实时增量/双轨迁移

1、割接主要流程：

- (1) 把最新备份数据在中间库恢复。
- (2) 把验证成功备份出来脚本在迁移目标 KES 数据库执行，完成结构迁移工作。
- (3) 使用 KDTS 工具把中间库的数据全量迁移到目标 KES 数据库，完成数据迁移。
- (4) 中间库和 KES 进行数据比对，验证无误。
- (5) 配置 KES 集群（如割接方案里有集群配置）。
- (6) 在既定割接上线时配置生产环境和 KES 的 KFS 双轨方案，实时同步增量数据。
- (7) 对生产库和 KES 进行数据比对，验证无误。
- (8) 切换应用访问 KES 数据库，支撑应用运行，并进行应用测试/调优。
- (9) 如一切正常，迁移成功。
- (10) 如发生意外，切换应用继续访问源数据库，待问题解决后再切换应用访问 KES 数据库。

四、割接上线方案

详细内容将在课件《割接方案和评审》章节培训。

07-备份容灾

详细内容将在课件《备份容灾》章节培训。

08-运维保障

详细内容将在课件《运维保障》章节培训。

KINGBASE®