

金仓数据库管理系统 V8.0

技术白皮书

版 权 声 明

©1999-2020 北京人大金仓信息技术股份有限公司 (Beijing Kingbase Information Technologies Inc.) 版权所有。

KingbaseES 是北京人大金仓信息技术股份有限公司和/或其分支机构的注册商标。本文中所涉及的其它商标或产品名称均为各自拥有者的商标或产品名称。

本产品白皮书中的信息如有更改，恕不另行通知。虽然已尽力确保本产品白皮书的完整性和准确性，但北京人大金仓信息技术股份有限公司对本产品白皮书的内容不作任何保证，包括任何暗示的保证。北京人大金仓信息技术股份有限公司对本技术白皮书中包含的错误或遗漏，或者因使用本技术白皮书引发的任何损失概不负责。

未经北京人大金仓信息技术股份有限公司许可，任何人或组织均不得以任何手段与形式对本技术白皮书内容进行复制或传播。

北京人大金仓信息技术股份有限公司

地址：北京市朝阳区容达路 7 号中国电科太极信息产业园 E 座 2 层

网址：www.kingbase.com.cn

客户服务邮箱：support@kingbase.com.cn

客户服务电话：400-601-1188

目录

1 KINGBASEES V8 概述.....	1
2 功能架构	2
3 支持广泛、可扩展的数据格式	2
3.1 内置丰富的常用数据类型.....	2
3.2 可扩展的类型系统.....	3
4 数据库应用程序开发	4
4.1 编程接口及标准符合性	4
4.2 应用开发框架	5
4.3 应用开发工具	5
5 最高可用性架构	6
5.1 处理计划外的故障.....	6
5.1.1 数据备份恢复	6
5.1.2 控制文件多副本	7
5.1.3 物理复制	7
5.1.4 逻辑复制	8
5.1.5 基于第三方 HA 软件的方案	8
5.2 最小化计划内停机时间	8
5.3 提前发现潜在风险.....	9
6 高性能处理各类应用场景.....	10
6.1 共性场景	10
6.1.1 数据装载	10
6.1.2 高效查询优化	11

6.1.3 缓存优化机制	11
6.1.4 数据分区方案	12
6.1.5 IO 校准	13
6.2 事务型应用场景	13
6.2.1 分布式集群	13
6.2.2 读写分离集群	15
6.2.3 共享存储集群	15
6.3 分析型应用场景	17
6.3.1 行列混存	17
6.3.2 GPU 插件	18
6.3.3 并行查询	18
6.3.4 物化视图	18
6.3.5 BRIN 索引	19
6.3.6 文本模糊查询	19
7 纵深防御，实现最高安全性	19
7.1 身份鉴别	20
7.2 多样化访问控制	20
7.3 用户数据保护	21
7.4 数据安全传输	21
7.5 安全审计与分析	22
8 低成本、快速进行数据集成	24
8.1 数据迁移	24
8.2 数据同步	24
8.3 数据集成	25
9 简单、高效进行系统管理	25
9.1 配置部署	25
9.1.1 集群配置	25
9.1.2 数据库管理	26

9.2 备份恢复	26
9.2.1 物理备份恢复	26
9.2.2 逻辑备份还原	26
9.3 运行管理	27
9.3.1 系统监控	27
9.3.2 SQL 日志分析	28
9.3.3 健康检查	28
10 完备的生态系统	29
10.1 上层生态支持	29
10.1.1 中间件	29
10.1.2 备份软件	29
10.1.3 商务智能软件	29
10.1.4 GIS 平台	29
10.2 底层生态（运行环境）支持	29
10.2.1 CPU	29
10.2.2 操作系统	29
10.2.3 云部署	30

1 KingbaseES V8 概述

KingbaseES 是北京人大金仓信息技术股份有限公司研发的数据库产品，单机支持不低于 50000 物理并发连接，单机单表可支持不低于 100TB 数据批量装载、数据插入、更新、查询、删除。兼顾各类数据分析类应用，可用做管理信息系统、业务及生产系统、决策支持系统、多维数据分析、全文检索、地理信息系统、图片搜索等的承载数据库。产品完全兼容主流数据库（如 Oracle、SQLServer、MySQL 等）语法的服务器端、客户端应用开发接口，以及索引、函数的创建执行。支持纯国产大型数据库，有安全、可靠的数据存储、备份方案，有近线、离线数据调阅方案，实现数据的网络存储及应急状态的数据本地存储。具有开放性，支持异种数据库的访问，包括实现对文件数据和桌面数据库的访问、实现对大型异种数据库的访问的能力等。

KingbaseES 汇集了人大金仓在数据库领域近二十年的技术积累，包括公司在国家“核高基”重大专项数据库课题的研究成果，是唯一入选国家自主创新产品目录的数据库产品，已广泛适用于电子政务、军工、电力、金融、电信、教育及交通等行业，是国家级、省部级项目中应用最广泛的国产数据库产品。

支持账本数据库，能够记录用户操作记录，当用户创建防篡改用户表时，能够通过数据库本身来记录用户表中每条数据的变更行为。支持查询用户表与其对应的历史表以及全局区块表中关于该表的记录是否一致。

作为 KingbaseES 产品系列最新一代版本，KingbaseES V8 在系统的可靠性、可用性、性能和兼容性等方面进行了重大改进，SQLbypass 对于简单 sql，跳过经典的执行器执行框架，并且直接调用存储接口，大大加速简单查询的执行速度

它包括以下主要技术特性：

➤ 高度容错，稳定可靠

针对企业级关键业务应用的可持续服务需求，KingbaseES V8 提供可在电力、金融、电信等核心业务系统中久经考验的容错功能体系，通过如数据备份、恢复、同步复制、多数据副本等高可用技术，确保数据库 7×24 小时不间断服务，实现 99.999% 的系统可用性。数据库内核自带检测坏块函数，可以通过该内置函数发现丢失的文件，并可以修复。对于表文件损坏的页面，能够通过备库自动修复，不借助任何工具。

➤ 应用迁移，简单高效

针对从异构数据库将应用迁移到 KingbaseES 的场景，KingbaseES V8 一方面通过智能便捷的数据迁移工具，实现无损、快速数据迁移；另一方面，KingbaseES V8 还提供高度符合标准（如 SQL、ODBC、JDBC 等）、并兼容主流数据库（如 Oracle、SQL Server、MySQL 等）

语法的服务器端、客户端应用开发接口，可最大限度地降低迁移成本。

➤ **人性设计，简单易用**

KingbaseES V8 版本提供了全新设计的集成开发环境（IDE）和集成管理平台，能有效降低数据库开发人员和管理人员的使用成本，提高开发和管理效率。

➤ **性能强劲，扩展性强**

针对企业业务增长带来的数据库并发处理压力，该版本提供了包括并行计算、索引覆盖等技术在内的多种性能优化手段，此外提供了基于读写分离的负载均衡技术，让企业能从容应对高负载大并发的业务。

注：下面为描述上的方便，把 KingbaseES V8 简写为 KingbaseES。

2 功能架构

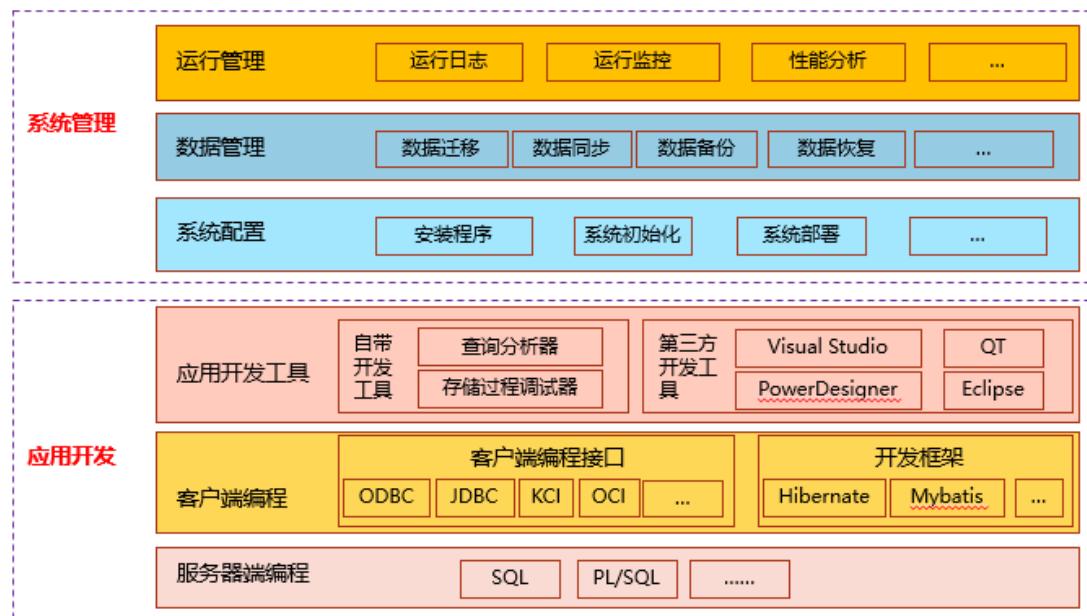


图 1 KingbaseES 逻辑功能架构示意图

3 支持广泛、可扩展的数据格式

3.1 内置丰富的常用数据类型

针对大部分应用常用的数据类型，支持多种 SQL 标准，支持 ANSI SQL-2003 以上标准；支持 JDBC/ODBC 标准接口；支持 PL/pgSQL 等过程语言。KingbaseES 提供了符合 SQL 标准和用户使用习惯的数据类型支持：

➤ **布尔**：支持符合 SQL 标准的布尔数据类型。

- 数值：支持符合 SQL 标准的 TINYINT、SMALLINT、INTEGER、BIGINT、NUMERIC、REAL、DOUBLE 数据类型，并兼容 Oracle 的 NUMBER。
- 字符串：支持符合 SQL 标准的 CHAR、VARCHAR、CLOB 数据类型，并兼容 Oracle 的 CHAR (CHAR|BYTE)、VARCHA。
- 位串：支持符合 SQL 标准的 BIT、VARBIT 数据类型；
- 二进制串：支持符合 SQL 标准的 BLOB 数据类型；
- 日期时间：支持符合 SQL 标准的 DATE、TIME、TIMESTAMP、INTERVAL 数据类型；
- XML：支持 XML 数据类型；
- 全文检索：支持用于对文本数据进行全文检索处理和查询的 TSVECTOR、TSQUERY 数据类型；
- 函数：包括：数值函数、分析函数、字符函数、类型转换函数、日期时间函数、聚合函数等，可以根据用户的实际需要定制特殊的函数。支持数值运算、比较运算、逻辑运算、字符拼接、类型强转等常用操作符。
- 高度兼容 Oracle，支持 number、date、blob/clob、varchar2、NVARCHAR2(N)、bfile 等数据类型，支持基于(+)操作符的表关联，层次查询 connect by 支持 level 伪列、支持 connect by isleaf,order siblings by 和 connect by root；
- 支持非关系数据操作：包括 SQL 对 XML 数据、地理信息数据、json 数据等多种数据的能力。
- 空间数据：支持符合 OpenGIS 规范的空间数据类型、函数、操作符，支持 oracle spatial 到 KingbaseES 的迁移。支持空间数据功能，支持丰富的原生几何数据类型，包括点、线、面、多点、多线、多面、几何集合等；支持 EWKT、EWKB 和 Canonical 格式的几何对象；支持 2D/3D 坐标系、坐标系转换 • 和球体长度计算；支持空间数据分析函数和聚合函数，包括 Area、Length、Distance、Extent、ST_3DLineInterpolatePoint 等；支持二元谓词如 Union 和 Difference，空间操作符如 Contains、Within、Overlaps、Touches 等
- 索引支持：支持多种类型索引，包括 B-Tree 索引、位图索引、函数索引、分区索引、Hash 索引、反向索引、字符串检索索引等，并支持在线重建索引。

3.2 可扩展的类型系统

在一些专业的应用领域，可能会有本领域专有的数据格式和数据处理需要，例如气象卫星领域的遥感遥测数据、物联网领域的传感器采集数据、图片搜索领域的模式和签名数据等。对于这类数据的存储和管理，通常是由本领域的应用开发人员通过自行编程的方式，在数据库外部进行管理。

针对这一类应用场景，KingbaseES 提供用户自定义数据类型的机制，允许用户根据自

己的业务要求定义专有数据类型、操作符和函数，从而一方面满足用户专有数据格式的存储和计算要求，另一方面实现跟其他常用数据类型的一体化管理。

4 数据库应用程序开发

KingbaseES 提供了符合国际标准的 SQL 语言及丰富多样的数据访问接口，支持与流行的集成开发环境紧密集成，并对主流数据库高度兼容，从而为应用程序开发者提供了便利。

4.1 编程接口及标准符合性

KingbaseES 为数据库应用开发人员提供了符合标准的服务器端和客户端编程接口，支持数据库监控，支持通过标准化数据库访问接口开发基于数据库的应用系统和软件产品。并提供多种开发语言和脚本编写语言的数据库访问支撑：

- SQL 标准：
 - 结构化查询语言 (SQL)：支持 SQL92、SQL99、SQL2003 标准，SQL92 入门级标准符合度达到 100%，SQL2003 核心级标准符合度达到 99% 以上，并兼容大部分 Oracle 常用语法。同时，支持与 Oracle 兼容的多种数据类型、索引、主外键约束、触发器、存储过程、包、匿名块、层次查询、视图、物化视图等。
- 索引推荐功能支持用户在数据库中直接进行操作，对用户输入的单条查询语句生成推荐的索引。
- 支持 AI 功能，通过结合深度强化学习和全局搜索算法等 AI 技术，实现在无需人工干预的情况下，获取最佳数据库参数配置。
- 开发接口与框架：
 - 兼容多种数据库开发接口与框架：包括 DPI、OCI、O040、OCCI、OOPI、OTL、QT、Pro*C、Python3、django、sqlalchemy、JDBC、hibernate、.NET DATA PROVIDER、EFDbProvider、EFCore、NHibernate、DDEX 插件、ODBC；
- 开发语言：
 - 支持多种开发语言：C/C++、Python、JAVA、.Net、Go、PHP、Node.js；
- 网络协议：
 - 兼容多种网络协议：包括 IPV4 协议、IPV6 协议、RDMA 协议。支持 IPv4 和 IPv6 双栈访问能力。支持多语言字符集，支持 UTF8、GB18030 等常用中文字符集，支持 Unicode 编码字符集，需兼容 MySQL 常用数据类型包括 Tinyint、Bit、Datetime、LONGTEXT、LONGBLOB 等，支持表字段 AUTO_INCREMENT 属性、支持 REPLACE INTO、SELECT INTO OUTFILE、DISTINCT 列可使用非 ORDER BY 列、SELECT 字段列表可以使用非 GROUP BY 列等语法；支持 CURDATE、DATE_ADD、DATE_FORMAT、DATE_SUB、EXTRACT、MAKEDATE、MONTH、QUARTER、

JSON_ARRAY、JSON_CONTAINS 函数；支持#符号进行注释，支持存储过程中使用 declare handler 语句；支持通过参数设置表名大小写、字段名大小写是否敏感。；

➤ 锁机制：

- 支持表锁和行锁级别的并发访问控制，提供良好的死锁处理机制；具有大规模并发处理能力。

4.2 应用开发框架

KingbaseES 支持 Hibernate、Mybatis 等常用 Java 应用开发框架，QtSQL、ADO 等常用 C++/C# 应用开发框架，允许用户选择应用程序开发框架，以帮助用户复用在现有开发资源和技能上的投资。

4.3 应用开发工具

KingbaseES 支持 Visual Studio .NET、Eclipse、NetBeans、JBuilder、PowerBuilder、Delphi、C++ Builder、JDeveloper 等流行的集成开发环境，以及 PowerDesigner、Visio、ERWin、ER/Studio 等数据库建模工具，从而保证适应广大开发人员的使用习惯。

KingbaseES 自身还提供了丰富、易用的应用开发辅助工具，可以帮助 KingbaseES 数据库的应用开发人员提高开发效率，有效管理开发过程。

➤ 对象管理工具：

应用开发过程中，开发人员需要经常操作数据库对象，若通过 SQL 语句实现，则需要熟悉掌握较多的 SQL 语句并准确使用，每个数据库除了提供兼容 SQL 标准的语句之外，也提供了自己扩展的丰富的语句，直接编写比较困难。KingbaseES 提供了对象管理工具，对于常用的数据库操作，包括显示对象定义、创建对象、修改对象、编辑一个表的数据、移动对象到其他模式、以某个对象为模板快速生成类似对象等，都提供了可视化且快捷的方式，无需开发人员编写 SQL 语句，可以大大提高应用的开发效率。服务器和客户端工具均支持简体中文和英文来显示输出结果和错误信息。

➤ 性能分析。

➤ 存储过程调试：

在开发过程中，对功能比较复杂的存储过程和函数，开发人员往往需要通过调试功能来辅助开发。KingbaseES 综合考虑开发人员的调试习惯，提供 PLSQL 调试功能，支持加载，设置断点，单步，进入，跳出等调试步骤，支持查看堆栈和当前变量等跟踪手段，为开发人员在调试使用时提供便利。

5 最高可用性架构

当前，几乎所有组织都依赖信息系统实现高效运营、快速分析信息及有效参与竞争。因此，组织的信息系统基础架构和数据库必须持续可用；否则，组织将会以收入损失、生产力损失和客户及信誉损失等结果为停机付出代价。

KingbaseES 通过采用最高可用性架构，为客户提供的一系列产品和方案能够解决无法预测的停机及计划停机造成的常见问题，并且能以最低的成本进行部署，集群在确保 RPO=0 的前提下，可以实现 RTO≤10s.。

5.1 处理计划外的故障

众所周知，IT 基础架构不可避免的会发生诸如服务器故障、磁盘崩溃或存储损坏、网络中断和人为错误等故障，这些故障可能会造成意外停机。KingbaseES 提供了从内部设计实现到现场部署的全方位的容错机制，可最大程度防止和减少意外故障造成的影响：

➤ 全面故障检查，第一时间发现错误

对所有可能发生故障的硬件设备，如硬盘、网络、CPU、内存等，KingbaseES 在使用前总是首先检查该资源的有效性，从而第一时间发现故障，避免故障蔓延和升级，导致无法预知的结果。

➤ 完备的系统冗余，消除单点故障

支持进程冗余、数据冗余等多种方式，以增加资源的方式换取可用性。支持数据备份，支持数据同步，支持控制文件多副本，支持备机只读，读写分离等多实例，保证数据和实例的冗余，避免单点故障。

➤ 自动监测处理故障，无需人工干预

KingbaseES 提供的各种高可用性架构，均提供故障自动检测、自动切换、自动回复能力能力，对于网络，心跳，进程等可能出现的异常都无须人为干预，实现故障的自动处理。

5.1.1 数据备份恢复

KingbaseES 具备完善的备份和恢复功能，支持备份恢复功能，支持逻辑备份、物理备份功能。逻辑备份支持按照多个级别的压缩比进行备份，支持按照指定字符集编码进行转储，支持按照指定条件转储指定部分的数据，支持加密、并行处理等功能。物理备份支持联机热备，支持全量备份、增量备份，并支持并行处理、zlib、pglz 压缩算法，支持按照多个级别的压缩比进行备份，支持进行备份集的合并，支持设置备份的留存策略，支持 PITR 指定时间点恢复。逻辑和物理备份均支持本地和远程备份。支持 B 树/堆表等级别的备份与还原。支持备份中断机制，可根据系统负载峰谷状态随时中止或恢复备份流程。支持单机与共享集群之间双向的备份恢复能力；。

（1）物理备份及恢复。物理备份与恢复对数据库的数据文件、日志文件、控制文件的

原始内容进行备份和恢复。其主要包括以下技术特性：

- 支持联机备份功能：联机备份可在不停机或不中断数据库业务处理的情况下，对数据库整个系统实施策略灵活的在线备份，其具备完全备份和增量备份两种方式。
- 支持并行备份：可大大缩短备份时间，减少备份操作对于应用系统的运行影响。
- 支持基于时间点的系统恢复技术：允许将数据库的状态恢复到当前运行时间前的某一个时间点，从而有效处理由于人为误操作带来的数据丢失。

（2）**逻辑备份及恢复。**逻辑备份通过 SQL 读取数据库中的元信息和数据，构建出各个对象的定义和数据，并写到备份文件中。逻辑还原则将备份文件中对象定义和数据还原至指定数据库。其主要包括以下技术特性：

- 支持全库、模式、表等多种粒度的备份和恢复。
- 支持并行，极大的提高了逻辑备份和恢复的性能。

（3）**集群备份恢复。**产品支持完善的备份恢复能力，可实现单机与共享集群之间的双向备份和恢复，保障集群层面的整体高可用和高可靠性。1GB 数据导入时间在 500 秒内，导出时间在 200 秒内。

（4）快速备份恢复。产品提供表数据快速备份功能，单节点单线程备份性能达到 700MB/秒以上。单节点 10 线程备份性能达到 1500MB/秒。1GB 数据完全备份时间应在 100 秒内，完全恢复时间应在 50 秒内。

5.1.2 控制文件多副本

控制文件是数据库系统中最为关键的文件。若发生服务器断电或者磁盘写操作错误等异常而引发控制文件损坏，会导致服务器无法启动无法提供服务，而这种情况下，恢复数据通常需要很长时间。KingbaseES 支持控制文件多副本技术，同时维护多个控制文件副本。若发现主控制文件损坏，用户可以使用同步更新的其他控制文件恢复主路径的控制文件，快速完成恢复，并不损失用户数据。

5.1.3 物理复制

KingbaseES 基于物理日志数据复制技术，通过在备机应用接收到的数据库日志的方式，实现与主机的数据同步。且只要主节点数据一产生日志，就会马上传递到备库，具有更低的同步延迟，因此也称其为流复制。

物理复制技术的主要特性如下：

- 基于物理日志复制的同步机制，确保主备机数据的一致性。
- 支持同步和异步两种复制模式，其中同步复制可以提供最大保护，主备节点之间数据实现完全一致，但对主机会产生一些性能影响；而异步复制，对主机没有性能影响，同一个时间点上主备机之前数据存在差异。两种模式为用户在效率和数据一致性中提供多种可选方案。

- 对于同步复制，KingbaseES 通过设置反馈级别和一致性级别两个配置项，可实现同时支持多个同步复制和多个异步复制，从而支持部署本地的强同步模式和异地容灾的强同步模式。
- 支持一主一备、一主多备和级联复制等多种物理部署方式，满足用户不同级别的容灾和可用性需求。
- 支持断点续传，如果主备节点之间的链接断开，可从断开点再次开始同步，数据不会丢失，不会重复。
- 支持自动、手动两种故障切换模式，可自动或按需进行服务切换。

5.1.4 逻辑复制

在实际应用中，有些场景物理复制无法更好的满足需求，如需要将多个数据库实例的数据，同步到一个目标数据库；或将一个数据库实例的不同数据，复制到不同的目标库；或在多个数据库实例之间，共享部分数据等场景，KingbaseES 可通过逻辑复制的技术予以支撑。

逻辑复制是对物理日志进行解析，通过将数据库日志中记录的对数据库的修改解析为 SQL，并在备库中执行这些 SQL 操作，从而实现主备节点的之间的数据同步。

逻辑复制技术具有如下主要特性：

- 适用于数据同步、数据上报、数据分发等多种应用场景；
- 支持一主一备、一主多备、级联复制、双向同步（双活）等多种部署方式。
- 支持行级过滤，变更的数据可以满足一定条件后再同步到备库。
- 支持断点续传，如果主备节点之间的链接断开，可从断开点再次开始同步，数据不会丢失，不会重复。
- 支持从 KingbaseES 向非 KingbaseES 的异构数据库进行数据同步；
- 支持插件式日志解析模块，可以方便的扩展为其他异构数据库的日志解析流。

5.1.5 基于第三方 HA 软件的方案

KingbaseES 还支持基于包括 Windows MSCS、Linux pacemaker 等常用第三方 HA 软件的高可用部署方案，包括基于共享存储和数据同步的多种部署方式。该方案可有助于用户构建一体化高可用方案，实现对信息系统各部件可用性的统一管理。

5.2 最小化计划内停机时间

通常硬件升级、软件升级和修补等基本维护是每个 IT 运营活动都可能涉及的内容，因此而造成的停机往往属于计划内停机。KingbaseES 提供了许多解决方案来帮助客户减少运维活动所需的计划停机时间，这些解决方案包括：

- **联机打补丁：**通过基于物理或逻辑复制技术构建的数据库集群，用户可以在联机状态下以“滚动”方式对集群节点应用数据库软件补丁：即可以方便地将用户从一台服务器迁移到另一台服务器；从集群中停止该台服务器，对其进行补丁修补，然后再使其

恢复联机。接着，对集群中的每台服务器重复同样的操作。此外，对许多一次性修复来说，可对KingbaseES实例直接应用补丁而不需要关闭该实例。

- **滚动版本升级：**使用基于物理或逻辑复制技术构建的数据库集群，实现备用数据库升级、测试新（升级后的）环境，然后将用户切换到新环境，无需任何停机。
- **联机重定义表结构：**在不间断支持联机生产系统的同时允许对表结构进行变更，从而减少维护停机时间。管理员可以允许最终用户对正在进行结构变化的表执行插入、更新和删除操作。
- **联机创建索引：**在不间断支持联机生产系统的同时对表创建索引，允许最终用户对正在进行结构变化的表执行插入、更新和删除操作。
- **在线增减节点：**在用户保持连接的情况下，轻松地为数据库集群添加或移除数据库服务器。

5.3 提前发现潜在风险

KingbaseES 提供多种风险检测手段，帮助用户提早发现数据库运行隐患，并且给出建议的解决方案，从而及早消除数据库的隐藏风险。

- **数据页面完整性校验。**KingbaseES 支持数据页面的完整性校验，写出或读入数据页面，系统会自动计算和校验页面的 CRC，若数据页面发生损坏，系统会及时发现，并进行隔离以防止错误蔓延。
- **健康状态检查。**KingbaseES 提供脱机和联机两种方式，对数据库系统的健康状态进行检查，包括数据库的存储情况、事务运行情况、REDO 日志的一致性、归档日志的情况、表空间的一致性、控制文件的一致性、数据页面的一致性、低于内存的使用情况等，进行全方位的检查，及时发现问题，以提前进行防范和解决；提供视图查看最新的以及历史上健康检查的结果。
- **运行状态可视化分析。**KingbaseES 提供基于数据库的动态指标，对数据库运行的事务、锁、查询、连接、检查点等信息进行全方位的分析，以可视化查看方式，供 DBA 对应用的运行情况进行分析，以达到应用优化，避免因数据库使用不当引发不必要的问题。通过分析服务器运行历史，可以及时展现服务器运行过程中可能出现的各类错误，以及系统运行的状态对比，如服务器数据的变化，数据库负载的变化等，便于用户实时跟踪查看系统运行中的问题。

6 高性能处理各类应用场景

随着企业业务数据量呈爆炸式增长，数据库性能问题变得越来越突出。通常，数据库性能优化是一个系统工程。而不同的场景下，对性能的关注点也会有所不同，KingbaseES 分别采用更有针对性的优化技术来保证用户的各种应用均可得到高性能的支撑。

6.1 共性场景

KingbaseES 具备主流成熟的数据库技术，提供数据批量处理技术，查询计划重用技术，结果集缓存技术，多版本并发控制技术，查询内并行处理技术，数据压缩技术，行列融合存储技术。支持查询计划的 HINT 功能。可支持不止一种的多版本并发控制（MVCC）机制，同一行的数据记录经过事务的修改，新老数据记录根据所选的机制不同，可存放在表的同一页面中，或者新数据存在表数据页中、老数据存放在专用的回滚空间中。多种 MVCC 机制可以更好的适应不同的业务场景。

6.1.1 数据装载

数据装载是使用数据库中非常常见的场景，而面对 TB 级以上数据的数据装载，性能问题最为凸显。数据库中最常用的 INSERT 语句，已经无法满足装载的性能需求，KingbaseES 提供了多种装载方式。

- COPY

支持 txt, csv 等多种输入格式，可以批量处理元组，比 INSERT 插入元组速度提高至少 1 个数量级。

- Bulk Load

通常 Copy 数据导入方式(即 Copy from 方式)的执行效率相对较低，在性能上无法很好满足海量数据的导入要求。针对这个问题，KingbaseES 提供了批量加载功能(即 Bulk Load 功能)。

在数据写入方面，批量加载支持以下两种方式：

- 1) DIRECT 方式

DIRECT 方式下，跳过共享缓冲区(shared buffers)缓冲数据和写 WAL 日志过程，直接把检查合格的导入数据写到数据表。比 COPY 方式，性能提升一倍。

- 2) BUFFER 方式

BUFFER 方式下，导入的数据使用共享缓冲区进行缓冲，并写入 WAL 日志。该方式支持并行，可以最大程度使用 CPU 和 IO 资源，在并行情况下性能可比 DIRECT 方式提升 5 倍以上。

- 自定义加载策略

支持自定义加载策略实现数据异步加载。能够满足数据合并和批量化要求。

- 并行加载

提供数据加载驱动程序和存储节点间并行加载支撑能力，数据装载性能能够随存储节点

和数和加载服务数据实现线性扩展，满足海量数据高校导入需求。

6.1.2 高效查询优化

查询优化是数据库系统的关键部件，它对于整个关系数据库的性能，尤其对于复杂 SQL 语句的执行性能至关重要。

在这方面，为满足多种应用场景的性能要求，KingbaseES 提供了多种查询优化技术，主要包括：

➤ 支持执行计划的手动干预，支持 Hint 功能（单表 Hint，多表 Hint，命令行 Hint）。

➤ 基于代价模型的查询优化

KingbaseES 采用基于代价的查询优化模型，它的基本思路是：对于一个给定查询，首先生成多个执行计划，然后依据系统收集的各种统计信息估算每个计划的代价。最后，查询优化程序自动按预定义的启发式规则从中选择次优或最优的执行计划。

需说明的是，该过程对于应用程序和最终用户来说是完全透明的。

➤ SQL 语句重写

KingbaseES 查询优化器可依据预先定义的启发式规则重写用户的 SQL 语句，它的基本实现思路是：依据预先定义的启发式规则，优化器将用户的 SQL 语句转换成可以返回相同结果、但执行效率更高的新 SQL 语句。新语句通过显著减少扫描、连接或聚合的次数达到提高 SQL 语句执行性能的目的。

➤ 多种类型的索引

KingbaseES 支持支持 B-TREE 索引、GIN 倒排索引、Gist 空间索引、UBtree 等多种索引访问方式。支持中文的全文检索功能，能够支持 like 操作符的全文检索，中文全文检索功能不依赖中文词典。还支持 BLOOM、RUM、BRIN、GIN、SP-GiST、GiST 等多种索引，分别使用与支持多个字段的任意字段组合查询、文本的模糊查询和数据快速加载等多种场景。支持唯一索引、表达式索引和部分索引。

➤ 支持 INDEX-ONLY SCAN

当索引列包含了所有查询结果目标列时，KingbaseES 可以使用 INDEX-ONLY SCAN，只扫描索引页面，就可以返回结果集，而无需再次扫描数据页面，减少 IO 消耗。

➤ 用户优先级控制

支持对用户优先级别的控制，可实现高优先级的用户执行的语句有限执行。

6.1.3 缓存优化机制

I/O 瓶颈一直是数据库性能的主要瓶颈之一，随着处理器性能的不断提高，这一问题变得更加突出。针对这个问题，KingbaseES V8 依据系统内部的统计信息把经常访问的用户数据放在内存，而把不经常访问的数据放在外存，以便在 SQL 语句执行过程中尽可能提升内存缓冲区的命中率，减少系统对磁盘的访问次数，从而有效应对 I/O 瓶颈。

➤ 支持系统表缓存大小自适应

KingbaseES 可根据服务器负载，自行动态调整系统表缓存大小，更合理的使用内存。当访问的表比较少时，可自动减少内存消耗；当访问的表较多时，可自动增加内存的使用，减少内外存的数据交换，从而提高性能。

➤ 支持数据缓存预加载

通常情况下，重启数据库之后，数据库缓存将被清空，当应用程序开始运行时，需要从硬盘读数据，因此在数据库服务器重启之后的一段时间，数据库的性能将受到一定影响。

KingbaseES 提供缓存预加载工具，可以将需要的数据预先加载到数据库服务区的缓存中，也支持加载到操作系统缓存中，这将解决以往数据库重启之后，需要先从硬盘读数据从而带来的性能影响。

6.1.4 数据分区方案

为提高系统性能和应对海量数据处理，KingbaseES 提供了水平分区解决方案——使用表继承的特性，提供水平分区方案。其实现原理是：依据实际应用的 I/O 访问特性，把表从横向或纵向划分为若干个子表，使得具有相同或相近 I/O 特性的数据存放在同一个子表。这样在用户访问数据库时不必访问被分区表的所有子表，而只需访问有限个子表，从而达到减少 I/O 访问的目的。

KingbaseES 提供数据分区方案，支持哈希、范围、列表、时间间隔分区类型，支持范围分区等常用类型；支持分区管理，分区表导入，分区创建，分区删除，分区 truncate；支持全局分区索引和局部分区索引，支持插入索引时选择索引所在表空间，支持索引重组。

数据分区方案包括以下优点：

➤ 多种分区方式

KingbaseES 为用户提供了多种分区方式，支持包括范围分区、哈希分区、列表分区，以及间隔分区等多种分区方式。支持多级分区机制。

➤ 提高可用性

DBA 对一个大表进行数据整理、数据备份等维护性操作时，一般要持续很长时间才能完成。在此期间，应用要么不能访问这些表，要么访问的速度会受到很大的影响。在采用分区技术后，对整个大表的单个操作被分解为对多个分区的操作，缩短了每个操作的时间和对并发访问的影响。在采用水平分区之后，单个分区的损坏不会影响其他分区上的查询和操作。

➤ 提高性能

KingbaseES 的水平分区通过剪枝优化，剪除 SQL 语句 where 条件中不相关的横向分区，这样应用中系统只需扫描剩余未剪除的横向分区，而不必扫描分区表的所有横向分区。在适用场景下，利用这些横向分区方法系统能获得指数级的性能提升。

➤ 扩展数据处理规模

KingbaseES 数据分区技术有效降低了系统的 I/O 访问压力，为 KingbaseES 采用分而治之的策略实现大规模海量数据处理提供了坚实的技术保障。

6.1.5 IO 校准

数据库服务器执行一个 SQL 语句会有多种可能的执行方法，服务器中的优化器会比较多 种执行方法的执行代价，选出代价最小的执行方法，从而达到优化的目的。对每一种执行方法进行代码预估时，IO 等硬件处理能力是比较重要的输入参数。因此，在不同的数据库安装环境下，为了输入准确的 IO 硬件处理能力，从而选择最后的执行方法，KingbaseES 提供了 IO 校准功能，能够根据随机生成的数据分析出具体硬件的处理能力，从而给出针对当前硬件环境下，数据库的最佳运行参数，引导运维人员对数据库进行调优。

6.2 事务型应用场景

数据库的使用场景可以分为事务性应用场景和分析性应用场景。事务性应用系统也叫 OLTP、操作型、事务处理型信息系统，主要用于业务流程电子化，面向的对象包括所有参与业务流程的人员，对数据库的操作特点是大量短、频、快的增删改操作。

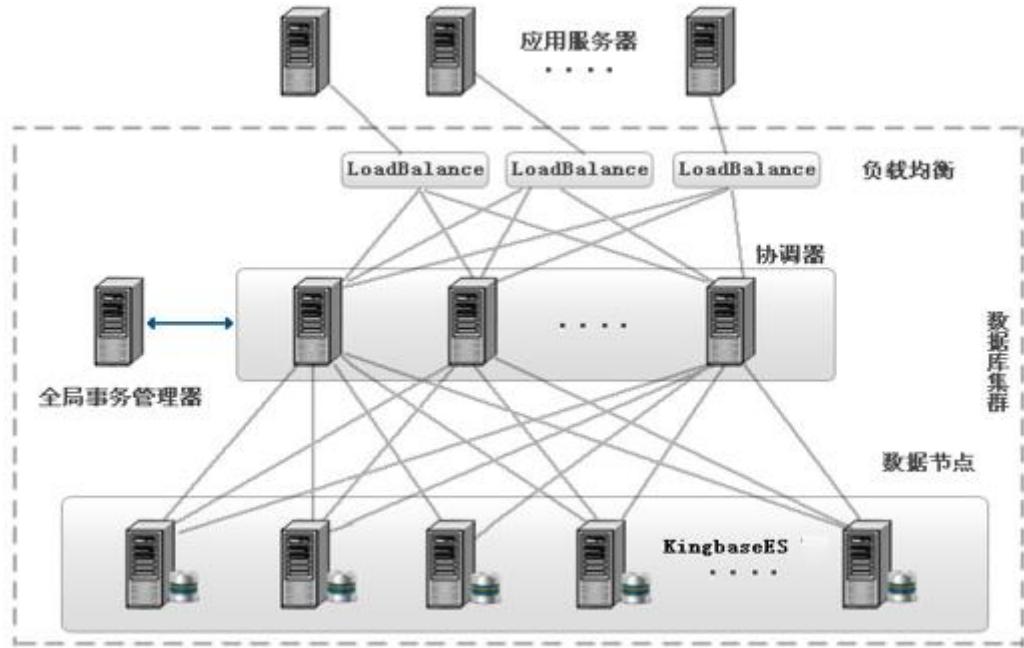
典型的系统包括企业系统信息，OA、ERP、CRM，包括业务支撑系统（BSS）、网络运营支撑系统（OSS）等。

在这种事务型应用场景中，除了使用上述共性的优化手段，KingbaseES 还提供了高效的并发机制，以及读写分离等多种手段提高性能。

6.2.1 分布式集群

KingbaseES V8 提供分布式集群部署能力，支持完全对等无共享（share-nothing）的 MPP 架构，支持数据多副本，支持的副本数不低于 7。通过分布式并行 SQL 处理实现处理并行化；利用多份数据副本机制来保证数据的高可靠性；并提供图形化的工具，监控和检测机制来加强系统的可靠性；通过在线节点扩展支持实现高扩展性；通过优化内存和在线数据压缩等措施，提升集群海量存储管理能力。

KingbaseES V8 提供集群数据副本功能，并支持副本的在线克隆、添加、删除。



同时, KingbaseES 支持透明分布式集群, 支持计算存储分离的系统架构, 实现计算、日志、存储三层分离, 可实现各层独立扩展、按需配置。既具有分布式数据库的高可扩展、高可用、高并发处理能力, 且对用户透明; 同时又具备传统数据库的所有高级特性, 支持传统数据库所有开发接口和业务开发框架的技术架构。

KingbaseES 服务器引擎利用高效的多线程共享体体系结构支持多用户并发。KingbaseES 采用多版本并发控制技术, 即一种 MVCC (Multi-version concurrency control) 与传统的两阶段封锁相结合的技术, 这种技术不仅读不阻塞, 并且读写操作也不相互阻塞, 从而显著地提高了事务的并发处理能力。过程性语言支持自治事务, 包括存储过程、自定义函数、触发器以及匿名块。存储过程支持自治事务的嵌套调用。

KingbaseES 集群具备精确查询秒级响应能力, 支持基于大数据智能索引的查询优化能力, 能够达到集群单节点单表基于时间列 (千亿级跨度) 精确查询秒级响应。

系统提供可扩展的逻辑和物理存储结构, 每个服务器可以支持多个单独运行的数据库实例, 每个实例可以管理多个独立的数据库, 单节点管理能力可达 10TB 以上。该技术通过支持由多台计算机组成的集群, 适应云计算环境下分布式海量数据处理场景, 将数据存储和处理并行分布化。

KingbaseES 支持非共享存储集群, 具有分布式并行计算能力, 支持分布式并行查询、支持并行数据加载。支持多种数据分部方式 (哈希表、循环表、复制表、单节点表)。能够对系统负载的实时监控, 实现动态并发控制, 有效提高资源的利用率。不影响系统运行的情况下在线添加、删除节点。

MVCC 机制使每个 SQL 语句看到的都只是一小段时间之前的数据快照（一个数据库版本），而不管底层数据的当前状态，从而可以保护语句不会看到可能由其他在相同数据行上执行更新的并发事务造成的不一致数据，为每一个数据库会话提供事务隔离。MVCC 避免了传统的数据库系统的锁定方法，将锁争夺最小化来优化多用户环境中的最佳性能。

KingbaseES V8 产品提供在线扩展功能，具有一定的在线扩展能力，即在不影响系统正常使用的情况下，增加数据库的数据出来能力。具备弹性扩展能力，支持在线添加存储节点，总的节点规模要求不低于 1024.

此外，KingbaseES 并发用户数目可依据用户实际需求进行配置，并可配的最大数目仅受限于系统资源和操作系统所允许的最大的进程数目。

6.2.2 读写分离集群

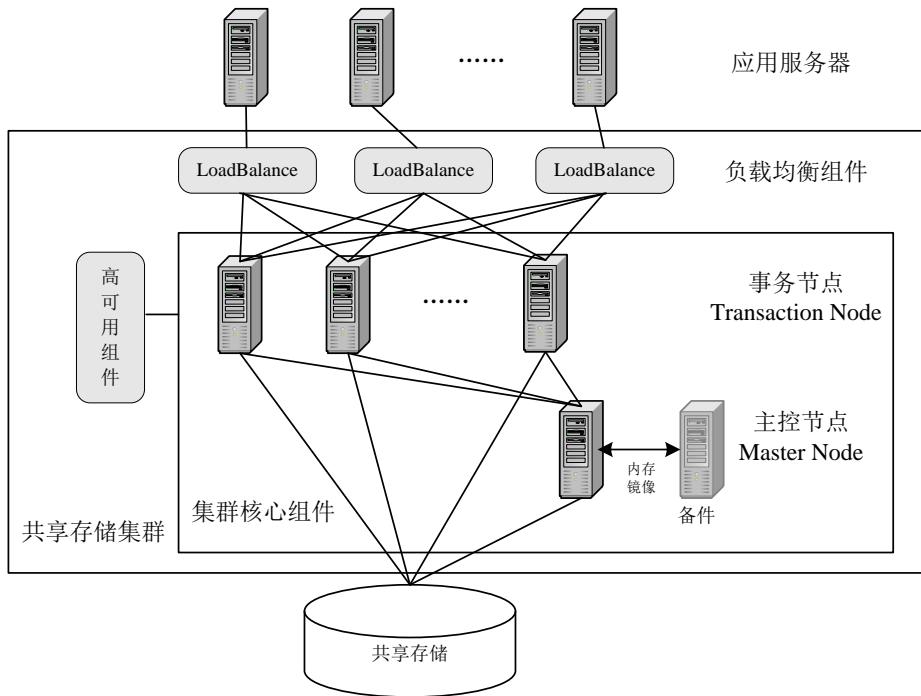
在很多应用场景中，读操作的数量远超过写操作的数量，针对这种读多写少的场景中 KingbaseES 推出了读写分离集群，可以大大提高整体的性能和可用性，读写分离对于读事务可以实现近 30%的线性加速比。支持读写分离以及故障转移功能，支持主从、一主多从、级联复制等架构，从库能设置延迟复制和优选提交复制。

KingbaseES 的读写分离中间集群支持 2 个或者更多的节点，在主节点中执行写事务，在从节点中执行读事务。主节点和从节点之间使用日志流进行数据复制。

当部署集群时，若需要配置集群负载均衡时，可设置 load Balance 参数。KingbaseEScluster 负责调度，后面的主节点和多个备节点执行用户请求。用户写请求发送主节点，只读请求发送给备节点。

6.2.3 共享存储集群

KingbaseES 支持高可用，支持共享存储集群，支持共享存储集群加远程实时备机容灾，远程实时备机支持单机和共享集群等多种模式。



如上图所示，整个集群分为以下组件：

- 集群核心组件
 - 主控节点 (Master Node, 缩写 MN)
 - 事务节点 (Transaction Node, 缩写 TN)
- 集群组件
 - 负载均衡
 - 高可用

各个组件的职责如下：

◆ **主控节点**

负责在内存中维护集群中的增量数据并对事务节点提供对这些数据的查询。

◆ **事务节点**

负责处理数据库服务请求，从共享存储的只读数据历史版本和主控节点上的内存数据版本合并出结果集返回。还负责根据主控节点上的内存数据更新共享存储的只读数据。

◆ **负载均衡组件**

负责统计各个节点的负载信息，监听和响应应用连接，并将连接按需转发到事务节点上。

◆ **高可用组件**

KingbaseES 监控主控节点/事务节点的软硬件故障、内部/外部网络故障、内部心跳网络故障、共享存储访问光纤网络故障，并保证在任何最苛刻的故障环境下集群系统的高可

用性。

共享存储集群包括以下主要技术特点：

◆ **高性能**

事务节点在提供服务时对共享存储只读，记录日志和更新只读历史版本都采用批量顺序写，屏蔽了随机写，可以充分利用设备的 I/O 能力，最大化 SSD 的优势；

事务节点采用无锁结构，减少冲突；

主控节点全内存操作，最小化响应时间。

◆ **高可用性**

KingbaseES 共享存储集群支持计算节点和存储节点分别部署在多个同城机房，实现存储、数据库服务的双活高可用。在 100KM 以上距离上实现了数据库服务、存储设备的高可用。

KingbaseES 支持处理各种节点，网络，存储访问光纤的故障处理，可以在只有一个节点可以正常工作的情况下仍然提供服务。

事务节点故障不影响系统提供服务，主控节点的快速恢复技术可以在故障时只影响短时间的系统服务。

◆ **高可扩展性**

事务节点支持 scale-out 和 scale-up，支持在线动态增减，可达到线性扩展；

主控节点支持 scale-up，可以通过虚拟化技术达成 scale-out。

6.3 分析型应用场景

分析性应用系统，主要用户统计分析和挖掘规律，面向的对象包括决策者和分析人员，对于数据库的典型操作包括大量数据入库，大量复杂查询分析等，典型使用场景包括决策支持系统、舆情分析系统、医疗统计系统、进行分析系统等。

6.3.1 行列混存

海量数据处理以 OLAP 分析型场景居多，这些场景中查询密集型应用的特点之一是查询一般只关心少数几个字段。但是，在数据库传统的行存储引擎中，即使只访问一个元组的少数组字段，也需把该元组的所有字段均取出，因此，这种设计局限在 OLAP 分析型应用中难免遭遇性能瓶颈。列存储就是针对上面问题给出的一个有效解决方案。列存储利用按列存储管理数据的方式，对特定查询只需操作少数组字段，而不必关心其他字段，这样大大减少读取的数据量并便于分析。当前，按列存储数据已被广泛认为可在分析型查询负载下获得更高的性能。内置内存引擎，是真正的存储引擎，可通过简单操作语句直接指定是否启用内存引擎，使用内存引擎的表上的性能比磁盘表性能提升 40% 以上。能够实现在同一个实例中内存表跟普通的磁盘表的共存，内存表支持 ACID、常用 SQL 语法、存储过程和数据持久化等功能

特性，针对内存表支持使用 PREPARE 语句的查询原生编译。

KingbaseES 提供了全新的按列存储和操作能力，支持行列融合存储技术：支持变更缓存、高级日志等功能实现列存高频插入效率的提升和行列自动同步机制。且行存储和列存储引擎可以同时工作，可以从容应对复杂多样的业务应用场景。其列存储的逻辑存储组织形式，不再像行存储那样，由表中的列组合成一条元组，元组是存取的最小单位，而是定义表中每列为存取的最小单位，查询只涉及表中的部分列时，其他列值将不被访问。在物理存储组织形式上，列存储将表中的数据按列存储在磁盘上，同一列的列值在磁盘上的位置是连续的，在很多场景中这可以带来 I/O 的效率提升。比如，把对列的聚集操作转化为对磁盘的顺序读，把对列的更新操作转化为对磁盘的顺序写。

6.3.2 GPU 插件

在处理器内核数量和 RAM 带宽上，GPU 有显著的优势，GPU 通常有成百上千的处理器内核，RAM 带宽也比 CPU 大几倍。

KingbaseES 可提供 GPU 插件，一方面利用 GPU 实现大并发的简单数值计算；一方面利用 CPU 完成复杂的处理，如磁盘 IO 的交互；并实现 CPU 和 GPU 并行处理，从而保证运算性能更为高效。在一些典型的分析场景下，使用 GPU 插件性能可以提升 10 倍以上。

6.3.3 并行查询

目前硬件环境多 CPU 很普遍，但是一个 SQL 语句在传统数据库中仍然只能使用 1 颗 CPU 的资源，CPU 资源无法充分利用；另一方面，数据库要处理的单个 SQL 的复杂程度却日益复杂。并行成为一个在现有硬件情况下提升服务器性能的最急迫的技术需求。

KingbaseES 支持并行计算框架，包括并行执行计划的代价估算，通过消息队列、数据交换和数据收集实现多个进程的数据交换以及并行执行框架。目前，顺序扫描，索引扫描，连接和聚集计算，以及排序等场景都可支持并行执行。

在实际应用中，KingbaseES 将依据系统负载情况，通过自动评估查询复杂度来判断是否启用并行查询，并能自动根据系统的负载情况确定并行度，从而避免由于并行执行导致对 CPU 资源的过度争用。

对大部分复杂查询语句，KingbaseES 可实现线性或近似线性的性能提升。它意味着：对同样的 SQL 语句，如果充分利用一倍的 CPU 资源，处理的响应时间就可以降低 50%。

6.3.4 物化视图

对于表连接或聚集等耗时较多的操作，可以提前计算并保存结果到物化视图中，在执行查询时，就可以快速得到结果，在有索引的支持下，效率更高。

另一方面，物化视图也可以提高数据的安全性，使得用户只能接触部分数据。

与普通视图相比，物化视图的数据需要不定时的刷新才能获取到最实时的数据。目前支持全部刷新和增量刷新，并支持并发刷新。

物化视图就像普通视图那样代表一个经常使用的查询结果集，只是结果集像一个普通

表那样存储在磁盘上。物化视图也可以添加索引，不像普通视图每次请求时重新生成，物化视图是及时的快照。它们只在特定时刻刷新。这可以极大地加快使用物化视图的查询的执行速度。无需在查询中使用普通视图或做复杂表关联或运行聚合函数，使用一个包含所需数据在磁盘的物化视图可以提高效率。

6.3.5 BRIN 索引

BRIN 索引，Block Range Index，简称 BRIN，这种类型的索引存储了表的连续数据块区间以及对应的数据取值范围，用于字段值和在表中的物理位置具有一定关联关系的大数据量访问。

一个典型的场景是流式日志数据，比如用户将大批量的数据按时间顺序不停的插入数据表，若要按照时间来访问这样的数据，以往需要创建 BTREE 紴引，支持范围查询或者精确匹配，但是 BTREE 紴引需要存储的信息量较大，如果用户数据量较大，索引将会占用较大存储空间。若使用 BRIN，索引可以变得很小。BRIN 可认为是全表扫描的切片，如果数据值分布和物理值分布的相关性很好，那么 BRIN 无疑是非常好的选择。

6.3.6 文本模糊查询

模糊查询是文本搜索领域常见的需求，主要包含前模糊（like %xyz）、后模糊（like %like%）、前后模糊（like %xyz%）、正则匹配。

对于前模糊和后模糊，KingbaseES 使用业界通用的 btree 紴引方式进行加速。而前后模糊、正则匹配这样的场景，KingbaseES 通过 trgm 技术来加快查询速度。

Kingbase 支持自有的 GIN、GIST 紹引，即在创建索引、查询字符串时都对内容进行三字节分词，然后通过索引查找关键字。对于精确查找，最终做一次条件过滤；反之，按照相似度从高到低顺序来返回结果。实测显示，在百万数据量的情况下，与全表扫描相比，KingbaseES 的 trgm 有 400–500 的性能提升。

7 纵深防御，实现最高安全性

在数据库安全方面，国外数据库厂商在我国销售的数据库产品的安全级别只能达到 TCSEC C1 或 C2 级，个别产品虽在 C2 安全级别基础上增加了部分如强制访问控制、数据加密等高等级的安全特性，但其整体安全级别均不能达到 B1 级或更高，从而不能很好满足如军队、军工、政府、金融行业、保险行业及电信行业等涉密部门的高安全要求。因此，研制具有自主产权的高等级数据库安全产品非常必要。

KingbaseES 自主研发高等级的数据库安全产品，它完全符合国家安全数据库标准 GB/T 20273-2006 的结构化保护级(即第四级)的技术要求，该级别近似等同于 TCSEC B2 级。在国产数据库厂家中，它率先通过公安部计算机信息系统安全产品质量监督检验中心的强制性安全认证，并获得销售许可证。

KingbaseES 通过全新的结构化系统设计和强化的多样化强制访问控制模型框架，自主

开发了多个高等级的安全特性，并完整实现包括多重身份鉴别、入侵检测与报警、可信路径、推理控制、形式化证明及隐蔽信道分析等在内的全部结构化保护级的技术和功能要求。

7.1 身份鉴别

KingbaseES 支持基于强化口令的身份鉴别，它包括对数据库用户施加口令复杂度检查、帐户和口令有效期限设置、帐户锁定等安全策略管理等机制。此外，KingbaseES 还支持基于 Kerberos、Radius、LDAP 认证协议和 CA 等技术在内的与第三方身份认证产品相结合的外部统一身份鉴别或集中化身份认证方式。

通过强化身份鉴别与 SSL 安全传输相结合，KingbaseES 建立了服务器与客户端的可信路径安全通道，实现了双向可信认证，杜绝假冒用户身份或者假冒数据库服务器的恶意行为，并进一步防范这些恶意行为可能引发的重放攻击。

7.2 多样化访问控制

如下图所示，KingbaseES 采用多种控制手段确保用户访问数据的合法性和安全性，并有效防止非法用户的未授权访问。

这些访问控制技术具体如下：

1. 自主访问控制

KingbaseES 采用 ACL(存取控制列表)技术实现了用户对于自有表和列(字段)的自主访问控制。用户可自主授权其他用户和角色对自有数据对象上的操作权限。当一个用户访问某个数据表/列时，自主访问控制将依据 ACL 检查该用户对访问对象的访问权限是否合法，用以决定接受还是拒绝本次用户的访问行为。

对用户和角色的存取权限的定义称为授权，这主要通过 GRANT 和 REVOKE 语句来实现。可授予用户的权限包括系统权限和对象权限，而对象权限包括表级权限和列级权限等。

2. 强制访问控制

KingbaseES 支持全密态等值查询功能，全密态数据库提供数据整个生命周期中的隐私保护，涵盖网络传输、数据存储以及数据运行态，密态等值查询通过技术手段实现数据库密文查询和计算，实现数据拥有者与数据管理者读取能力分离。支持动态数据脱敏，在不改变源数据的前提下，通过在脱敏策略上配置针对的用户场景 (FILTER)、指定的敏感列标签 (LABEL) 和对应的脱敏方式 (MASKING FUNCTION) 来灵活地进行隐私数据保护。按照 Bell-La Padula 安全保密模型，并基于实用性和灵活性的考虑对强制存取模型策略进行了扩充。所谓强制访问控制是指主体用户对数据库客体对象访问的安全策略是由系统强制实施的，任何用户都无法突破系统定义的强制安全策略而非法访问未授权的数据库客体对象。与自主访问控制不同的是，强制访问控制由安全管理员管理而不是数据对象属主，并它直接在受控主客体上施加安全标记，而不再从 ACL 中获取权限信息。

除了支持简单保密性原则—“从下读、平行写”外, KingbaseES 还可以支持“区间写”。为了支持一些特殊的管理性工作, 系统还提供了特权功能。拥有特权的主体能绕过相应的标记仲裁规则。

3. 推理控制

在数据库系统中, 访问控制限制了用户对敏感信息的访问, 即只允许用户访问授权信息, 而禁止用户访问未授权信息。这种方式虽然防止了敏感信息的直接泄露, 但却无法防止这些信息的间接泄露, 例如:

假设某个公司信息系统中, 姓名和职务属于授权信息, 而工资属于未授权信息。那么, 用户 A 有权限访问姓名和职务信息, 而无权访问工资信息。此外, 还假设用户 B 职务和 A 相同, 存在函数依赖职务 (授权) \rightarrow 工资 (未授权), A 的工资是 1000 元 (假设每个用户都知道自己的工资), 则用户 A 可推出 B 的工资也是 1000 元。上述这类问题无法通过传统的访问控制方式解决, 需借助推理控制技术。

所谓推理问题是绕过访问控制, 授权用户利用已知敏感信息和推理方式获取非授权信息的一种攻击方法。而恶意用户获取非授权信息的途径称为推理通道。什么是推理控制? 所谓推理控制就是对推理导致敏感信息泄露问题的一种对抗方法。

7.3 用户数据保护

用户数据是数据库系统的核心资源, 其安全性对整个系统至关重要。通常, 用户数据中包含具有保密要求的重要信息, 如商业秘密信息。一旦这些信息被窃取或篡改, 将给用户造成极大损失, 甚至带来灾难性的后果。因此, 保障用户数据安全是非常必要的。

KingbaseES 为用户数据提供多方位的保护, 有效杜绝对恶意用户对存储数据的非法窃取, 及外部破坏和篡改, 这些保障措施具体如下:

1. 数据存储完整性保护

通过在每个数据块头增加“数据水印”, KingbaseES 实现了数据存储过程中的完整性校验和保护。每次读磁盘时, 自动进行数据水印校验。每次写磁盘时, 自动更新数据水印。这样, 可以有效杜绝来自数据库外操作系统层的非法篡改。整个过程对用户完全透明。

2. 数据备份安全保护

备份数据是否安全可靠一直是数据库安全中容易被忽视的技术点和管理点。针对这个问题, KingbaseES 提供带安全标记和带密钥的加密备份和还原, 同时施加完整性数据水印, 且支持备份压缩, 可充分保证了数据备份和恢复的安全性和可信性。

7.4 数据安全传输

传统的万维网协议 HTTP 不具备安全传输机制。它采用明文的形式传输数据、不能验证通信双方的身份、无法防止传输的数据被篡改等, 从而无法满足电子商务和网上银行等应用

的安全性要求。

在信息传输方面, KingbaseES 采用成熟的 SSL 技术, 支持通过 SSL 协议实现客户端和服务器之间的安全数据传输, 使得数据在传输过程中难以被窃听、篡改、重放和伪造。在此基础上, 还支持使用国家保密局审批的第三方 SSL/IPSec VPN/加密机产品来为企业级用户提供硬件级的高性能网络通信安全保证。

7.5 安全审计与分析

任何系统的安全保护措施都不是完美无缺的, 蓄意盗窃、破坏数据的人总是想方设法打破控制。而审计是应对这种情况的一种常用安全保障措施。审计功能将用户对数据库的所有操作自动记录下来并放入审计日志中。事后, 审计管理员可以通过对审计日志的分析, 对潜在的威胁提前采取有效的措施加以防范。

在这方面, KingbaseES 配置了独立的安全审计系统, 设置了专门用于存储审计数据的系统表和专门的管理员角色, 并对数据操作和用户行为进行全面监控与审计。审计管理员可定义与数据库安全事件相关的审计规则, 并能对指定用户和数据表, 以及对身份鉴别、自主访问控制、标记、强制访问控制等相关操作进行审计。审计管理员通过相关的审计视图可查询和分析审计记录。

实际应用中, KingbaseES 提供不同级别的审计配置策略, 提供按照用户、语句类型等维度进行审计; 可对审计范围进行阈值设置。能够记录用户操作记录, 当用户创建防篡改用户表时, 能够通过数据库本身来记录用户表中每条数据的变更行为。支持查询用户表与其对应的历史表以及全局区块表中关于该表的记录是否一致。能满足系统对不同级别安全级审计的要求, 支持会话级日志跟踪, 支持对会话所执行的 SQL 语句, CPU 耗时, 执行计划, 使用内存等内容进行记录系统日志。

如果当前登录用户的操作在审计规则的审计范围内, 则将触发系统生成一条审计记录。此外, 出于易用性考虑, KingbaseES 还提供一套与 Oracle 兼容的审计视图。

◆ 审计类型

KingbaseES 支持多种类型的审计功能, 可实现细粒度审计:

(1) 语句审计

对数据库中执行的 SQL 语句进行审计。语句审计只针对语句本身, 而不针对语句所操作的对象和用户。如: 利用 AUDIT ALTER TABLE BY SCOTT 语句, 可以对用户 SCOTT 执行过的所有 ALTER TABLE 语句进行审计。

(2) 特权审计

对数据库中的系统权限使用情况进行审计。如利用 AUDIT CREATEROLE 语句, 可以对所有利用 CREATE ROLE 进行用户或角色权限管理的操作进行审计。

(3) 对象审计

对某个指定对象(表、视图、序列、存储过程、函数等)的某一类操作进行审计。如利用 AUDIT ALTER ON EMPLOYEES 语句, 可以对 EMPLOYEES 表中所有的 ALTER 操作进行审计。

(4) 用户审计

对数据库系统中的数据库用户进行审计。如 AUDIT SELECT BY SYSDBA WHENEVER SUCCESSFUL, 则数据库管理员 SYSDBA 执行的所有成功查询操作将被审计。

(5) 策略审计

对强制访问控制的安全策略相关操作, 如策略的创建、删除、生效等进行审计。

◆ 预置审计模版

一直以来数据审计的配置和管理对于安全管理员来说是一个很高难度的工作, 除了要求有很高的安全防护意识外, 还需要对于数据库审计的功能特性和管理接口要有系统清晰的掌握, 这在安全数据库系统的架设初期是个很艰巨的任务。

为了提升数据库审计的可用性和易用性, KingbaseES 从登录会话、系统权限、管理特权、语句行为等方面, 在安权衡全性和性能基础上, 预置了高、中、低三套审计模版。审计管理员可以很方便地阅读、学习和套用模版, 定制适合用户安全应用场景的审计规则并保存为自定义模版。此外, 审计模版还支持导入和导出, 能够在其他 KingbaseES 系统上复用。

◆ 审计记录保护

审计是安全数据库的最后一道屏障, 审计系统必须绝对保证审计记录的可信性和安全性。KingbaseES 对审计记录提供了多种安全保护措施, 充分保证审计系统的可用性和审计记录可信度。

KingbaseES 中, 只有审计管理员可以查阅审计记录, 且包括审计管理员在内的所有用户均不允许对审计记录进行插入和修改。审计管理员无法修改和删除审计记录, 只能通过转储方式腾空审计记录空间。转储的审计记录在被自动加密和施加完整性保护后, 形成内部保留审计记录, 保留审计记录不允许清除。

而且, 随着审计记录不断增长, 磁盘空间难免出现不足, 这就不可避免地要考虑如何规避防止审计记录溢出和审计服务不可用的问题。KingbaseES 通过可配置的预警和动作机制保证了审计服务的可用性。

◆ 审计分析

KingbaseES 提供了专门的审计记录分析功能, 通过定制和应用相关的审计过滤规则, 系统会自动筛选生成与规则相关的审计记录。然后, 通过对筛选记录统计分析, 可以为数据库系统的安全状况提供量化依据。

8 低成本、快速进行数据集成

近几十时间随着信息化的推进，我们可以步入了大数据的时代。数据量庞大。数据格式和质量参差不齐，严重阻碍了数据在各部门和各软件系统中的流动与共享。因此，如何解决数据的互通共享，如何对数据进行有效的集成管理已成为数据管理领域中一个非常突出的问题。

8.1 数据迁移

KingbaseES 提供了功能强大的数据迁移工具，支持 B/S 部署模式，支持库级、模式级、表级数据的导入导出，允许设置不同的用户、角色，可多人同时登录执行迁移相关工作。迁移工具需具备评估、自动转换、迁移、数据校验等核心能力，至少支持 Oracle、MySQL、SQL SERVER、DB2、JDBC、Kafka 作为数据源；支持对数据库结构信息、存储过程以及应用 SQL 进行兼容性评估分析，生成的评估报告，可预估迁移人工改造工作量、可以提供 SQL 语句的改造建议、支持通过扫描 MyBatis 的 mapper 文件进行兼容性评估；支持数据库对象 DDL 自动转换、包括存储过程和自定义函数的自动改写、编码转换、特殊字符自动处理等；迁移完毕后，可针对表验包括数据行数和数据内容的一致性；迁移工具需提供全图形化操作界面，在评估或迁移任务有进度条和百分比显示、提供图形界面设置单批次数据迁移抽取的行的数量、并行度、自定义迁移规则、集中批量的迁移规则定义为迁移模板；迁移工具需支持全量迁移、增量迁移。同时支持目标数据库到源数据库的反向增量迁移；支持在线手工改写。可以通过 JDBC 连接处理包括 Oracle(8i、9i、10g、11g)、纯文本文件等数据源的数据。可实现 KingbaseES 之间以及 KingbaseES 与异构数据库之间的数据迁移和结构迁移。

KingbaseES 提供主流数据库迁移，sql 脚本、文件迁移等功能，支持图形化向导式完成迁移工作。迁移工具支持并行化数据迁移、批量数据快速加载等方法，支持为每个迁移对象制定独立的迁移策略，支持配置提交批大小、索引迁移选项等，并允许将迁移策略应用到批量的对象。对于迁移过程中出现的异常情况，迁移工具能够记录并保存异常信息。

可支持数据类型的自动转换，并保留了源数据库中对象的属性(例如 Oracle 约束的启用/禁用状态、存储过程的有效/失效状态)，极大地减少了系统移植的工作量。

此外，KingbaseES V8 的数据迁移工具还提供了缓存迁移模式和命令行模式，经过配置后的数据迁移工具可以实现自动化且高效的数据迁移任务。

8.2 数据同步

KingbaseES 提供物理同步和逻辑同步等多种方式，可以完成同构和异构数据库的持续同步。

物理同步支持数据库级同步，支持实时同步、异步同步和级联同步，KingbaseES 支持同时配置超过 1 个的实时同步节点，可以达到金融级数据同步要求。

逻辑同步则更加灵活，支持异构数据库，可以选择同步部分对象的部分数据，包括同步某些模式，默写表的满足指定条件的数据，同步的备机同时可以执行其他更新操作。

支持 OLTP 数据库、OLAP 数据库的混合集群式部署，通过数据库内部数据同步机制实现数据信息在 OLTP 数据库引擎和 OLAP 数据库引擎之间的快速流转。

8.3 数据集成

KingbaseES 提供了多种数据集成手段，包括：

- 支持多媒体数据类型、空间数据类型；
- 使用 json，集成半结构化数据；
- 支持同构数据库跨库操作；
- 支持外部表，集成多种外部数据源，包括异构数据库，各种文件等，支持直接读取和更新操作；
- 支持与 RDS 的集成，直接访问 RDS 中的数据；
- 支持与大数据平台、kafka 等数据进行集成。

通过以上集成手段，KingbaseES 可以把不同来源、不同格式、不同特点性质的数据在逻辑上有机地集中和处理，帮助用户实现更全面的数据共享和融合。

9 简单、高效进行系统管理

KingbaseES 提供了强大的数据库自动管理功能，能够根据数据库当前的运行状态实现数据库性能参数的自动调整。支持多种部署方式，支持分布式、集中式部署。配置、管理、优化工作大部分由系统自动完成，减少了 DBA 的工作负担。系统也提供了易用的图形和命令行管理工具集，供具有不同经验的系统管理员和应用开发人员使用。

KingbaseES 数据库管理工具使 DBA 和应用软件开发人员能够有效地管理多个 KingbaseES 数据库系统。管理工具可以跨平台管理 KingbaseES 的多个安装实例。管理工具具有统一的图形化或命令行界面，功能强大，界面友好，操作方便。

9.1 配置部署

9.1.1 集群配置

KingbaseES 提供全功能的数据库双机热备能力，完全支持单机一致的功能，不依赖于第三方集群组件，支持备机可读。可在不停机状态下在单机系统与主备系统间平滑变换。备机时刻保持与主机的数据同步。一旦主机发生故障，备机立刻可以自动切换成新的主机，继续提供服务，同时支持客户端透明切换。

KingbaseES 集群部署工具使用引导式的方式一步步引导用户完成集群的部署和配置，除此之外，KingbaseES 集群部署工具还支持如下操作：

- 集群运行状态监控
- 负载均衡节点的操作（启动、停止）
- 后台节点的操作（启动、停止、手动切换、加入）

9.1.2 数据库管理

KingbaseES 数据库管理工具用于数据库的管理和维护，提供简便快捷可视化的操作。

其中包括：

- 对象管理：对数据库、模式、表、索引和视图等数据库对象进行管理；
- 程序管理：对触发器、函数、存储过程等 PLSQL 相关内容进行管理；
- 存储管理：对数据文件，表空间和日志文件的存储管理；
- 支持表压缩功能，能够分别支持行存表压缩和列存表压缩。
- 公共对象管理：对各种数据类型和所有系统函数的维护和管理；
- 安全管理：对数据库用户和角色，以及权限管理；
- 日志分析：对数据库运行日志进行分析，图表化的展示数据库运行状态；
- 会话管理：对当前数据库会话进行查看和管理；
- 参数管理：对数据库配置参数进行查看和修改。

9.2 备份恢复

KingbaseES 系统提供多种备份还原能力，支持逻辑备份恢复、跨平台恢复、物理备份恢复、在线离线备份、增量级完全备份还原、基于时间点的还原、备份的并行处理及压缩加密。

包括：全库加密、备份文件加密、日志压缩加密等。另外，还提供恢复功能和备份校验功能，提供备份集功能，可对备份文件进行查看、删除管理操作，支持对数据文件进行块恢复。

9.2.1 物理备份恢复

数据库备份是运维人员日常工作中很重要的一部分。KingbaseES 提供了图形化的物理备份恢复工具，以应对数据库系统级的联机/脱机备份恢复工作。让运维人员轻松面对因发生事务故障、系统故障和介质故障等灾难事故对业务数据可能造成的损害。

物理备份工具支持：

- 支持恢复至系统崩溃前的任何时刻，且保证系统数据的一致性
- 支持图形化，一键式备份和恢复；
- 支持命令行，便于自动化和周期化。

9.2.2 逻辑备份还原

数据库的逻辑备份还原提供了多种粒度的对象的备份还原操作，方便用户根据各种情况制定备份恢复策略。若数据库中只有部分表更新频繁，那么可以及时备份更新频繁的表，而修改较少的对象，跨度时间较长执行一次备份即可，这样处理备份还原时间更短，占用外存

存储空间也会较小。

KingbaseES 逻辑备份还原工具提供了数据库对象一级的联机备份还原功能，提供模式对象、表及其相关约束，支持制定列分隔符、行分隔符和转义字符。

备份对象包括：

- 数据库
- 模式
- 表
- 视图
- 约束
- 权限
- 触发器
- 函数
- 序列

逻辑备份的输出格式包括：

- 二进制
- SQL 脚本

此外，针对表数据，还支持多种格式的导出：

- CSV
- HTML
- JSON
- XML

从而帮助用户可灵活调整、实施数据的备份计划。

9.3 运行管理

KingbaseES V8 支持集中式运维管理工具，包含数据迁移工具、性能监视工具等功能，同时提供集群部署功能、监控功能和告警功能。支持监控主机运行状态 (CPU 利用率, 内存, I/O 等) 和远程主机上数据库的运行状态(事务, I/O, 会话等)，并且能够保存历史数据，通过对数据的分析提供相关的性能分析报告

9.3.1 系统监控

数据库性能和运行状态是运维人员平时工作指标的一部分，KingbaseES 提供数据库系统的监控与跟踪能力，可以动态监视数据库的进程、锁、存储以及 SQL 操作等执行状态，帮助 DBA 很好地跟踪数据的历史情况当前情况，使 DBA 对于数据库的运行状态有清晰的了解，从而有效地帮助 DBA 找出系统的性能瓶颈。

同时，系统提供丰富的系统表和系统视图，可使维护人员快速获取系统元数据信息，包

括但不限于：集群节点信息、集群节点状态信息、表信息、列信息、表数据节点分部规则信息、用户信息。

KingbaseES 监控工具支持的系统监控内容主要有：

- SQL 日志筛选、记录。
- 当前数据库中活动的进程；
- 存储使用的历史情况；
- 数据库中存在的锁；
- 数据库中死锁的历史信息；
- 当前数据库中缓存区的使用情况；
- 数据库中 SQL 语句执行的统计信息。

9.3.2 SQL 日志分析

KingbaseES 提供 SQL 日志记录功能：支持异步日志，SQL 日志的筛选记录机制，可根据 SQL 类型、SQL 耗时等素组合指定日志策略，帮助系统运维人员对重点 SQL 进行监控。多日志文件支持，支持为不同用户设置不同文件，降低运维人员对日志的分析难度；

KingbaseES SQL 日志分析主要内容有：

- 支持异步日志；
- 支持 SQL 日志筛选；
- 支持 SQL 类型、耗时查看；
- 支持对数据库运行日志进行分析；
- 图表化的展示数据库运行状态；

9.3.3 健康检查

数据库支持数据库自适应管理，支持自动存储管理、自适应内存管理、自调优管理，支持内场自调优、存储自调优、执行计划自调优。

数据库服务器数据的健康状态是数据库运维的最重要的指标之一，KingbaseES 健康检查工具提供全方位的数据库健康状态检查，包括：

- 文件完整性
- 页面完整性
- 逻辑完整性
- 日志完整性
- 归档完整性
- 数据字典完整性
- 版本一致性

对于检查出的异常情况，KingbaseES 健康检查工具会提供具体的修复建议。

系统提供 SQL 查询分析器, 支持 SQL 输入辅助功能, 包含了 SQL 语法检查功能和 SQL 输入助手功能。SQL 语法检查功能对用户输入的 SQL 语句进行实时的语法检查, 定位错误的 SQL 语法。SQL 输入助手能够对用户输入 SQL 进行实时的智能提示, 提示的内容包括数据库对象和 SQL 关键字等。管理工具支持提示机制, 可以向开发和运维人员提示当前操作界面所连接的数据库系统类型为生产库, 或是开发/测试库;

10 完备的生态系统

10.1 上层生态支持

10.1.1 中间件

KingbaseES 支持东方通、金蝶、中创、普元、华宇等厂商的中间件软件。

10.1.2 备份软件

KingbaseES 支持迪思杰、爱数、鼎甲、上海英方、火星高科等厂商的备份软件。

10.1.3 商务智能软件

KingbaseES 支持永洪、亿信华辰、SAP、润乾软件、皕杰软件等厂商的商务智能软件。

10.1.4 GIS 平台

KingbaseES 支持超图、中地时空、吉威数源、ArcGIS 等厂商的地理信息系统。

10.2 底层生态（运行环境）支持

10.2.1 CPU

KingbaseES 支持多核 CPU 服务环境, 支持 X86、X86_64、Itanium、SPARC、PowerPC 系列, 同时指出龙芯、飞腾、申威、兆芯、华为等国产 CPU 硬件体系结构。

10.2.2 操作系统

KingbaseES 所支持的操作系统包括:

- 微软 Windows 7、Windows XP、Windows 2003、Windows 2008 等 32 位及 64 位发行版本
- 各种主流的 Linux 操作系统 32 位及 64 位发行版本
- 中标麒麟、银河麒麟操作系统
- 中科方德操作系统
- 凝思磐石安全操作系统
- 一铭服务器操作系统
- AIX、Solaris、HP-UX 等 UNIX 操作系统

这些不同操作系统版本的服务器和管理工具之间具备无缝的互操作能力。

10.2.3 云部署

KingbaseES数据库支持云部署方式，同时也提供了云环境中的一体化管理工具。目前已经实现与阿里云、太极、华为、华三、世纪互联等的云平台的适配。

数据库和相关工具，提供 docker 镜像和相关解决方案，支持云化快速部署。

金仓数据库目前提供两种 docker 镜像，并可基于 docker 形式的自动配置部署：

1. 单机数据库系统
2. 读写分离集群数据库系统