



# 实现 GraphDB 数十倍性能加速， 快手利用 KVDK 释放英特尔® 傲腾™ 持久内存性能潜力

## 摘要

业务量的不断提升，以及数据规模的快速增长凸显了企业在数据处理方面的巨大挑战。为了高效处理海量数据，为用户画像、个性化推荐、风控、知识图谱、社交网络等应用提供复杂型数据关系运算能力的支撑，越来越多的企业开始强化在图数据库（Graph Database，GraphDB）方面的创新与投入。在提供了更高速的数据处理与分析性能的同时，GraphDB 给存储等关键组件的性能带来了巨大的挑战，传统 NVMe 固态硬盘已经出现了性能瓶颈，无法充分满足大规模 GraphDB 的应用需求。

为了给业务的快速增长提供高性能数据库能力的支撑，快手推出了基于开源技术优化的图数据库应用：快手 GraphDB，并通过英特尔® 傲腾™ 持久内存来构建高性能、大容量的存储资源池。快手 GraphDB 还使用了专为英特尔® 傲腾™ 持久内存设计的键值存储开发套件（Key-Value Development Kit，KVDK），以充分释放英特尔® 傲腾™ 持久内存的性能优势，并为数据库工作负载提供高级功能。

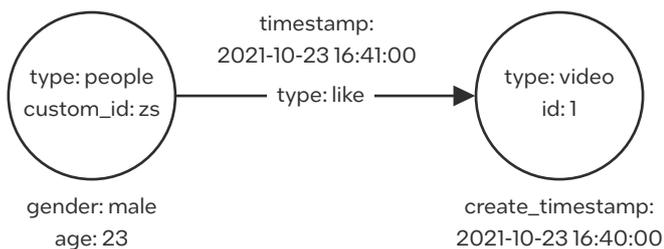
## 背景：GraphDB 在快手中的应用以及性能瓶颈

传统的关系型数据库能够应对大部分场景的需求，但是随着业务数据量和深度的增加，企业增加了在处理复杂数据关系运算方面的需求。例如，在社交网络、推荐系统、数据分析等业务中，用户的属性往往能够形成非常大的实体图，数据量往往会达到百亿级别。关系型数据库无法在有效的时间内计算出结果。所以，为了更好地体现数据间的连接，企业需要一种将关系信息存储为实体、灵活拓展数据模型的数据库技术，这项技术就是 GraphDB。

GraphDB 一般由顶点 (Vertex)、边 (Edge)、属性 (Property) 等关键数据来描述。其中，顶点表示一个实体，比如一个人/视频/设备；边表示顶点之间的关系，比如一个用户关注其他人，设备属于一个人等；属性表示除顶点 (或边) 之外的更多信息。

GraphDB 将图形的显式建模和查询平滑地集成到标准数据库环境中。对于标准应用程序，它提供了一些面向对象建模的关键特性，例如组织成层次结构的对象类、对象标识和引用对象的属性。另一方面，该模型允许通过将对象类划分为简单类、链接类和路径类来显示表示图，这些类的对象可以被视为图的节点、边和显式存储的路径。

与传统关系型数据库相比，GraphDB 能很好地体现并处理数据之间的关联关系，其从底层适配了图模型的数据结构，使得它的数据查询与分析速度更快。GraphDB 还具备非常灵活的数据模型，用户可以根据业务变化随时调整图数据结构模型，可以任意添加或删除顶点、边，扩充或者缩小图模型等。此外，GraphDB 符合当今最流行的敏捷开发需求，便于快速迭代<sup>1</sup>。



KGraph	Graph	Type	Vertex/Edge	Property
Relational	Database	Table	Row	Column

图 1. 快手 GraphDB 图数据特征

尽管具备众多优势，但是 GraphDB 依然具备诸多瓶颈，这包括：In-memory Graph 解决方案恢复时间长，无法适应大型数据集；线性扩展性有待提升，以动态地满足业务发展需求；在性能方面，GraphDB 物理节点的性能往往需要超过 2000 万 QPS，单顶点查询的延迟更低，传统 NVMe 固态硬盘会出现性能瓶颈，无法更好地提升 GraphDB 的性能。受单节点性能的限制，传统的 GraphDB 要想支撑更大规模的数据，往往需要通过扩展节点的方式来实现，但是这种方式无疑会带来较高的服务器采购、部署、运维等相关成本，在经济性方面并非最优选择。

因此，快手希望能够寻找到面向 GraphDB 应用的创新存储介质与存储引擎，尽可能地提升单节点的 GraphDB 吞吐性能，降低延时，解决在扩展性、恢复时间等方面的瓶颈，同时更好地控制 GraphDB 应用的总体拥有成本 (TCO)。

## 解决方案：通过英特尔® 傲腾™ 持久内存 (PMem) 与 KVDK 优化 GraphDB

为了化解 GraphDB 在性能、延时等方面的瓶颈，快手与英特尔合作，将 GraphDB 系统底层的存储介质替换为英特尔® 傲腾™ 持久内存 (PMem)，并采用 KVDK 作为 GraphDB 的存储引擎，以替代之前使用的 RocksDB。新旧方案的架构对比如图 2 所示。

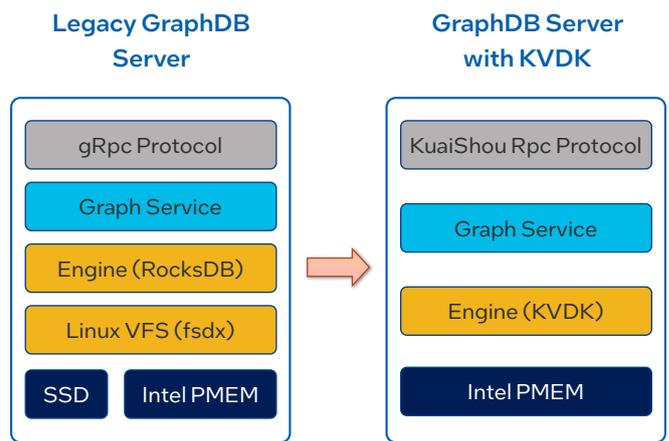


图 2. 快手 GraphDB 新旧方案架构对比

<sup>1</sup> RH Güting, Iv P.I. GraphDB: A Data Model and Query Language for Graphs in Databases[J]. proc.int.conf.on very large data bases, 1995.

与传统 NVMe 固态硬盘相比，英特尔® 傲腾™ 持久内存拥有巨大的性能优势。英特尔® 傲腾™ 持久内存是一种颠覆传统的持久内存，集高速、高性价比、大容量、持久数据保护和高级加密等优势于一体，并提供了内存模式 (Memory Mode) 和应用直接访问模式 (App Direct Mode) 两种模式。在内存模式中，它提供有 128 GB、256 GB 和 512 GB 容量，与普通的易失性系统存储器完全一样，但成本更低，能在稳定的系统预算中实现更高容量。

在应用直接访问模式下，应用程序和操作系统会有两种可用的内存。系统将决定是从 DRAM 还是英特尔® 傲腾™ 持久内存中写入或读取数据。要求最低延迟且不需要永久数据存储的操作可以在 DRAM 上执行，必须长期保留的大型数据结构和数据将由英特尔® 傲腾™ 持久内存进行存储。在此模式下，即使系统关闭电源，数据仍保留在内存中。

为了充分发挥英特尔® 傲腾™ 持久内存的性能优势，并为数据库工作负载提供高级功能，快手使用了 KVDK 来代替 RocksDB。KVDK 是一个用 C++ 语言实现的 key-value 存储库，它专为英特尔® 傲腾™ 持久内存而设计，并为易失性和持久性场景提供统一的 API，有助于提升英特尔® 傲腾™ 持久内存的性能表现。除了提供 key-value 存储的基本 API 外，它还提供了一些高级功能，如事务、快照等。

KVDK 搭载了持久内存分配器，能够避免分配线程的争用问题，提高核心可扩展性。根据英特尔® 傲腾™ 持久内存的特性，该分配器对于数据的顺序访问进行了优化，从而更好地提升了 RocksDB 的性能表现。此外，分配器还对于英特尔® 傲腾™ 持久内存的访问机制进行了调整，从而节省了持久内存的带宽。

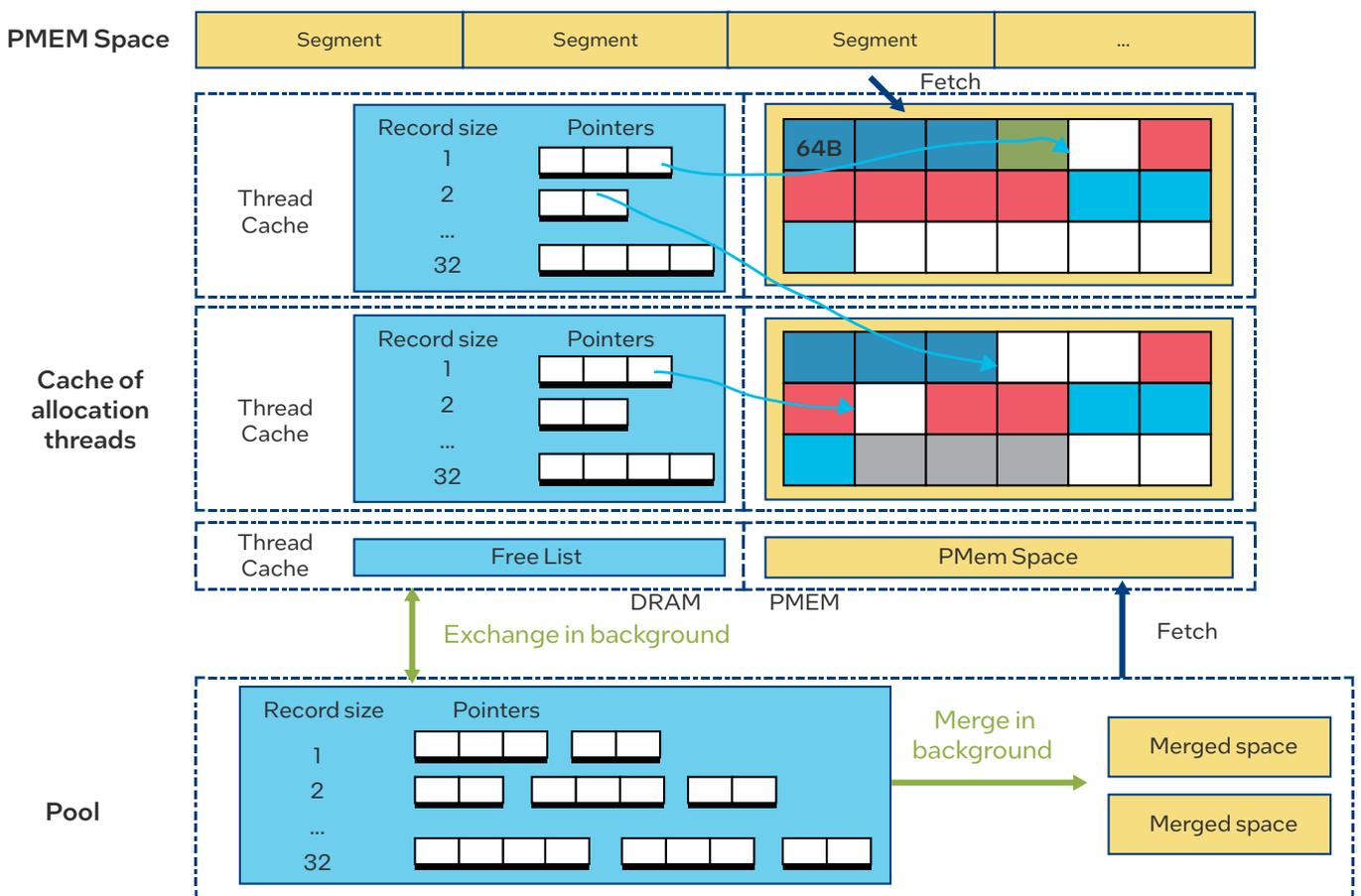


图 3. KVDK 持久内存分配器架构

KVDK 还通过哈希 (Hash) 算法进行未排序索引，并通过 Hash + 跳表 (Skiplist) 来实现排序索引，具备 CPU 缓存友好、无锁读、细粒度写锁、更少且可预测的持久内存开销、无锁读/寻等优势，能够更好地提升性能水平，并降低资源开销。

为了进一步发挥 KVDK 的优势，快手在 GraphDB 构建中删除了多余的 Linux 文件系统，并开发了一个新的更快的 RPC 框架以符合要求。

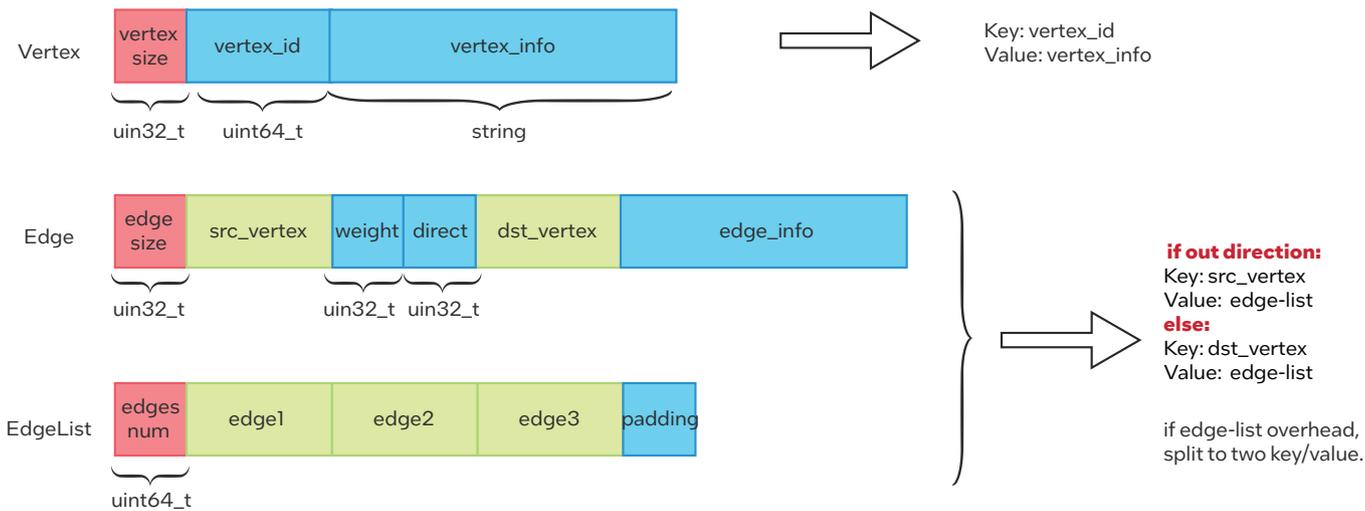


图 4. 快手在 GraphDB 中进行了针对性优化

### 收益：数十倍性能提升

为了验证基于 KVDK 的快手 GraphDB 方案带来的性能提升，快手进行了三个项目的测试。在第一个测试中，测试数据集为 30,000,000 个顶点 + 1200,000,000 个边，测试随机搜索构建数据的耗时；在第二个项目中，测试数据集为 1,000,000 个顶点 + 100,000,000 条边，通过构建热搜数据 + 4 degree bfs 搜索，测试延迟表现；第三个项目则测试从 Top N 个结果中计算出 Top 10 的时间。

测试数据显示，在三个项目中，基于 KVDK 的快手 GraphDB 在性能上都显著超过基于 RocksDB 的历史方案，性能领先幅度最高可达近 26 倍。

通过采用英特尔® 傲腾™ 持久内存以及 KVDK，快手显著优化了 GraphDB 的性能，能够充分满足百亿级别图数据库对于性能的要求，加快搜索、计算的速度，为用户画像、个性化推荐、知识图谱、社交网络等应用提供复杂型数据关系运算能力，支撑了快手创新型应用的高效运行。

同时，由于单节点性能的提升，因此在同等性能需求下，快手只需要消耗更少的服务器资源，就能让 GraphDB 系统达到设计的性能目标，节约了由此带来的硬件采购、空间占用、能耗、运维等方面的支出，提升了平台的经济性。

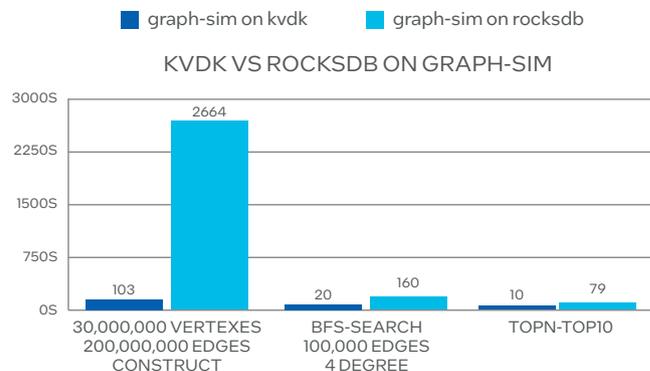


图 5. 快手 GraphDB 在替换 KVDK 前后的测试数据

“在 GraphDB 的性能优化过程中，选择高性能的存储介质，以及实现了针对性优化的存储引擎至关重要。通过与英特尔合作，并利用 KVDK 存储库，我们充分发挥了英特尔® 傲腾™ 持久内存的性能优势，并化解了在扩展性、恢复时间等方面的瓶颈，从而构建了领先的 GraphDB 系统，为多种应用提供复杂型数据关系运算能力，并为后续的深度应用奠定基础。”

— 刘鑫  
快手

## 展望：英特尔® 傲腾™ 持久内存赋能数字化创新应用

作为一种介于内存与存储之间的创新介质，英特尔® 傲腾™ 持久内存已经被证明在 IO 密集型负载中有着巨大的价值。在数据库应用中，其不仅可以替代内存型数据库的 DRAM 内存，提供大容量、经济型的内存集，也可以作为 GraphDB 等数据库的高性能存储介质，提供远超传统固态硬盘的性能表现。结合 KVDK 等存储库，能够更进一步发挥英特尔® 傲腾™ 持久内存的优势，为数据库提供强大的基础资源。

快手在 GraphDB 的实践证明，通过面向英特尔® 傲腾™ 持久内存的软硬件协同优化，能够提升数据库的性能表现。双方计划在未来强化合作，进一步推动 IT 基础架构领域的创新，增强 IT 基础设施在性能、扩展性、稳定性、可用性等方面的表现，为快手蓬勃发展的互联网创新业务构建坚实底座。

### 关于快手

快手诞生于 2011 年 3 月，最初是一款用来制作、分享 GIF 图片的手机应用。2012 年 11 月，快手从纯粹的工具应用转型为短视频社区，用于用户记录和分享生产、生活。如今，快手已成为大家记录和分享生活的平台，截至 2022 年一季度，快手应用的平均日活跃用户达 3.46 亿。

### 关于英特尔

英特尔 (NASDAQ: INTC) 作为行业引领者，创造改变世界的技术，推动全球进步并让生活丰富多彩。在摩尔定律的启迪下，我们不断致力于推进半导体设计与制造，帮助我们的客户应对最重大的挑战。通过将智能融入云、网络、边缘和各种计算设备，我们释放数据潜能，助力商业和社会变得更美好。如需了解英特尔创新的更多信息，请访问英特尔中国新闻中心 [newsroom.intel.cn](http://newsroom.intel.cn) 以及官方网站 [intel.cn](http://intel.cn)。



实际性能受使用情况、配置和其他因素的差异影响。更多信息请见 [www.Intel.com/PerformanceIndex](http://www.Intel.com/PerformanceIndex)

没有任何产品或组件是绝对安全的。

英特尔技术可能需要启用硬件、软件或激活服务。

具体成本和结果可能不同。

英特尔并不控制或审计第三方数据。请您审查该内容，咨询其他来源，并确认提及数据是否准确。

© 英特尔公司版权所有。英特尔、英特尔标识以及其他英特尔商标是英特尔公司或其子公司在美国和/或其他国家的商标。