

## 基于东方国信企业版 Hadoop 和英特尔® 傲腾™ 持久内存的大数据平台联合解决方案

### BONC 东方国信

“在测试中，英特尔傲腾持久内存对于大数据平台企业版 BEH 的提升作用非常明显，其在金融等行业中的部署有望帮助用户显著提升即时数据处理应用的性能，同时能够为用户降低 IT 成本，简化基础设施，我们还将在更多的应用场景中对英特尔傲腾持久内存进行测试，在此基础上构建新一代数据分析和人工智能应用平台的最佳解决方案。”

— 王虎

东方国信技术中心总经理

我们正处于一个数据爆炸式增长的时代，大数据正在成为企业、社会和国家层面重要的战略资源。全球科技行业着眼于以大数据驱动增长和创新，深挖数据富矿，从而提高工作效率，改善生活质量，最终带动经济增值。对于企业而言，通过对海量数据进行近即时的处理与分析，可深入洞察消费者需求，快速做出明智决策，获得重要商业价值。

作为由 Apache 基金会所开发的一种分布式系统基础架构，Hadoop 支持用户在不了解分布式底层细节的情况下，开发分布式程序，充分利用集群进行高速运算和存储。在开源 Hadoop 的基础上，北京东方国信科技股份有限公司（以下简称：东方国信）对 Hadoop 组件进行增强和封装，推出了适用于超大规模数据存储和在线分析的东方国信大数据平台企业版（BEH）通用产品解决方案，可以帮助企业快速搭建大数据平台，提升开发效率，降低后期维护成本。

为了在大数据分析应用中进一步提升 BEH 的性能表现，东方国信在服务器节点中使用了第二代英特尔® 至强® 可扩展处理器，并使用高性能的英特尔® 傲腾™ 持久内存作为 Spark SQL 的缓存设备，以加速缓存性能。测试结果显示，无论是在单节点还是集群测试环境中，英特尔傲腾持久内存都帮助 BEH 方案实现数倍的性能提升，以及显著的成本节约。

### BONC

#### 客户痛点：

- AI、实时大数据分析等应用需要更强大的性能支撑
- 满足目标性能的系统建设成本高昂

#### 东方国信 大数据平台 企业版 BEH

#### 为何选择英特尔傲腾持久内存：

- 加速系统缓存，避免性能瓶颈
- 优化总体拥有成本（TCO）

#### 价值定位：

- 延迟更低/响应更快/性价比优势更明显

## 面向大数据分析的一站式东方国信大数据平台 企业版 BEH

东方国信大数据平台企业版 BEH 在原生 Hadoop 基础上，面向最终企业客户进行了一系列的企业级功能增强，支持运维人员开箱即用，在多环境（物理和云化环境）下实现一键部署，统一管理和运维。同时，开发人员可使用工具化编排功能简化数据分析、加工过程，满足大数据时代海量数据在线分析应用的迫切需求。

与原生 Hadoop 相比，东方国信大数据平台企业版 BEH 具备以下优势：

- **开发编排：**支持批量/流式/非结构化数据的采集、计算和存储的可视化编排与开发，降低大数据使用门槛，提升使用大数据平台的能力。
- **部署运维：**支持一键式快速按需定制构建大数据基础平台，集中式/可视化平台管理，向导式操作完成运维管理，降低人员技术要求，减少误操作几率，保障平台安全。
- **组件适配：**适配多个社区版本、多种操作系统、多种服务器，实现 HBase/HDFS/Spark 等原生组件企业级应用部分功能增强，支持异构混合部署场景。

- **云化弹性：**支持基于 Kubernetes+Docker 容器的大数据组件弹性服务，支持部分组件云边结合，比如：基于 Kafka MirrorMaker 实现两级边云集群数据复制。
- **安全权限：**提供统一的大数据安全及权限管理，支持集群参数的可视化配置和滚动升级，支持可视化多租户资源隔离共享和可视化细粒度的权限访问控制。

东方国信大数据平台企业版 BEH 针对企业的实际需求进行了大量的科研攻关和技术创新，可以有效降低各种大数据处理应用的技术门槛和人才门槛，为各类企业和政府部门提供了开展海量数据在线分析相关业务的能力，将有力推动中国大数据相关应用的落地和发展。

### 解决方案：基于英特尔® 傲腾™ 持久内存优化 缓存性能

企业数据正在快速增长，同时越来越多的业务需要从实时数据分析中获得洞察，这意味着对于大数据分析性能的追求不会到达终点。例如，在面向上亿用户的实时推荐系统或面向全网的 RTB 广告竞价投放系统中，通常要求存储上亿级用户数据，同时在进行实时推荐或 RTB 广告询价等业务。此类应用对并发和时延要求极为苛刻，通常需要单服务器每秒支持上万次数据存取操作，对于性能提出了较为严苛的要求。

 <p><b>BEH-Developer（开发）— 加强批流结合一体化</b></p> <ul style="list-style-type: none"> <li>• 提供流式数据采集/计算的可视化的大数据开发和编排，帮助客户提升大数据开发效率。</li> </ul>	 <p><b>BEH-Security（安全及权限）— 用户资源权限访问一体化</b></p> <ul style="list-style-type: none"> <li>• 提供大数据访问权限及资源、计算、细粒度权限控制，帮助用户提升数据安全。</li> </ul>
 <p><b>BEH-DMA（部署及运维管理）— 加快物理容器一体化</b></p> <ul style="list-style-type: none"> <li>• 提供大数据集群管理和运维功能，帮助客户降低后期的维护成本。</li> </ul>	
 <p><b>BEH-Component（组件）— 多版本多服务器一体化</b></p> <ul style="list-style-type: none"> <li>• 基于开源组件进行多组件集成和部分组件源代码增强，为客户提供统一的大数据的存储和计算能力。</li> </ul>	
 <p><b>BEH-on Cloud（云化）— 拓展云化云边化</b></p> <ul style="list-style-type: none"> <li>• 基于 k8s + docker 虚拟化容器技术，实现大数据组件云化，帮助客户提供简化部署和弹性伸缩服务。</li> </ul>	

图 1. 东方国信大数据平台企业版 BEH

## 第二代英特尔至强可扩展处理器

- 为企业数据中心提供高可扩展性
- 与上一代处理器相比，为虚拟基础设施提供更高性能
- 实现卓越的资源利用效率和敏捷性
- 针对数据中心解决方案，提高数据和工作负载的完整性，确保监管合规性

东方国信大数据平台企业版 BEH 推荐采用第二代英特尔至强可扩展处理器。



东方国信大数据平台企业版 BEH 推荐使用基于第二代英特尔至强可扩展处理器的服务器。第二代英特尔至强可扩展处理器专为数据中心现代化革新而设计，能够提高各种基础设施、企业应用及技术计算应用的运行效率，进而改善总体拥有成本（TCO），提升用户生产力。它拥有更高的每核性能，能够在计算、存储和网络应用中，为计算密集型工作负载提供高性能和可扩展性。同时，得益于英特尔® 超级通道互联（Intel® UPI）、英特尔® Infrastructure Management（Intel® IMT）技术、英特尔® 高级矢量扩展指令集 512（Intel® AVX-512）等领先功能，它可满足严苛的 I/O 密集型工作负载的需求，能够帮助大数据应用着力打造性能更强的敏捷服务和突破性功能。

此外，第二代英特尔至强可扩展处理器内置人工智能加速，并已针对工作负载进行优化，提供真实应用场景中由客户驱动的性能，能够为各种高性能计算工作负载、AI 应用以及高密度基础设施带来一流的性能和前所未有的内存带宽。同时，采用矢量神经网络指令（VNNI）的英特尔® 深度学习加速（英特尔® DL Boost）显著提高了人工智能推理的表现，与上一代产品相

## 英特尔傲腾持久内存可以提供

- 更低的总体拥有成本（TCO）：要组建性能或容量相当的内存池，将仅需使用更低的成本，这也有助于企业降低基础设施的 TCO，获得更大的竞争优势。
- 敏捷的应用模式：企业用户可以根据自己的应用需求，在不对硬件进行调整的前提下，自由地使用英特尔傲腾持久内存的内存模式或是应用直接访问模式。
- 强大的性能，远超传统固态硬盘。

东方国信大数据平台企业版 BEH 推荐采用英特尔傲腾持久内存。



比，性能提升高达 14 倍，这为相关的人工智能应用奠定了坚实的基础。

I/O 是 Spark SQL 查询的常见痛点，即席查询语句的性能经常被受限于 I/O 性能。在应用实例中，东方国信发现，在混合并发查询等场景中，由于需要访问的表数据较多，内存很快消耗殆尽，东方国信大数据平台企业版 BEH 常常需要将大量数据缓存到相对低速的磁盘中，导致出现显著的磁盘 I/O 阻塞，且 CPU 利用率会上升至高达 100%，严重拖慢了系统速度。

为了进一步提升性能表现，同时有效控制总体拥有成本（TCO），东方国信采用了英特尔傲腾持久内存 + DRAM 作为缓存。除了相对于 DRAM 存储器的成本优势之外，英特尔傲腾持久内存还带来了内存子系统架构的颠覆性变化，将类似于 DRAM 存储器的字节寻址能力和类似于存储的持久性合二为一。这种结合意味着它可以直接映射到应用程序地址空间，消除了与传统存储的读写相关的瓶颈，使其成为工作数据和长期存储的主数据层。

英特尔傲腾持久内存可提供融合高速、高性价比、大容量、持久数据保护、高级加密等优势于一体的内存选项。不同于传统的 DRAM 内存，它集大容量、经济性和持久性于一身，读写带宽比 NVMe 设备高数倍，并大幅降低了延时问题。英特尔傲腾持久内存提供了两种不同的操作模式：内存模式（Memory Mode）和应用直接访问模式（App Direct Mode）。在内存模式中，它与普通的易失性（非持久性）系统存储器完全一样，但成本更低，能在保持系统预算的同时实现更高容量。模块的容量有 128 GB、256 GB 和 512 GB 几个版本。

在东方国信大数据平台企业版 BEH 应用中，英特尔傲腾持久内存作为应用直接访问模式的形态存在。通过启用应用直接访问模式，用户能够以更低的成本扩展一体机内存池容量，在服务器中同时使用独立的易失和持久性内存存储，加快使用基于传统 DRAM 的数据缓存结构的操作速度，进一步提高性能。

### 测试：消除磁盘 IO 阻塞，提升数倍性能

测试人员选取和东方国信大数据平台企业版 BEH 业务场景比较相近的 TCP-DS 标准 SQL 查询，评估使用英特尔傲腾持久内存 Cache 的 OAP 带来的性能提升，环境配置如表 1 所示。其中，OAP 是基于 Spark SQL 开发的项目，提供了包括缓存和索引的功能，用于加速 Spark SQL，可以满足秒级甚至更高要求的即席查询需求。

CPU 信息	2* 英特尔至强金牌 6248 处理器
物理内存信息	4* DDR4 64 GB 4* 英特尔傲腾持久内存 128 GB
组件版本	apache-hadoop-2.8.5 spark-2.4.4.patched (Intel 编译发行) JDK1.8
nodemanager 内存资源	DRAM 170 GB
测试程序	OAP-TPCDS-TOOL
磁盘	1*HDD
操作系统和内核	CentOS Linux release 7.4.1708 (Core) / 3.10.0-693 GNU/Linux

表 1：测试环境配置

首先，测试人员通过调整英特尔傲腾持久内存的初始化内存大小，观察 SQL 查询效率的提升，测试数据如图 2 所示：

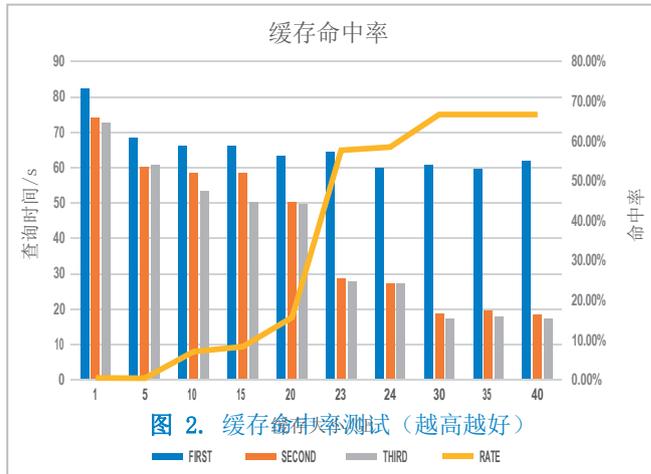


图 2. 缓存命中率测试 (越高越好)

测试结果显示，在初始化的缓存空间较小的时候，缓存并没有给第二/三次查询带来明显的效率提升，而且命中率显示几乎是 0%，这是由于缓存空间较小时 Cache 会发生缓存内容替换导致。此外，命中率和缓存空间并不存在线性关系。当缓存容量增大至 30 GB 之后，这些缓存足可以保证几乎所有的数据都缓存到英特尔傲腾持久内存中，继续增加缓存并不会给查询速度与命中率带来明显变化。

随后，测试人员模拟生产随机不同查询 SQL 时，英特尔傲腾持久内存对查询性能的提升效果。本组测试分为 Parquet only、Parquet + 英特尔傲腾持久内存 cache、Parquet + DRAM cache 三种模式（其中包含缓存的两组模式设置 4 种缓存大小规模，分别为 1 GB/12 GB/30 GB/60 GB），每组测试顺序执行 9 个 IO 较高的 SQL，共执行三次，测试数据如图 3 所示：

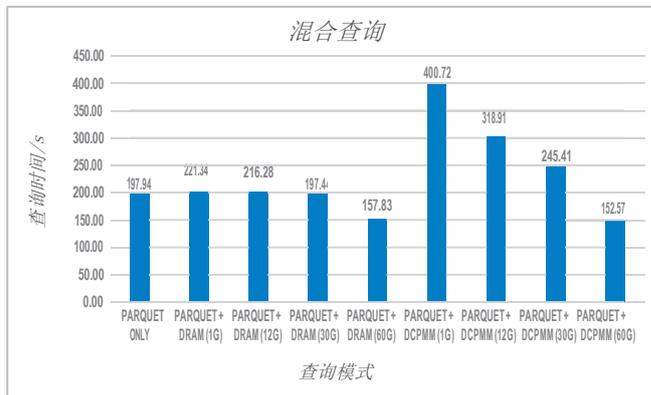


图 3. 混合查询执行时间测试 (单位：秒，越低越好)

测试显示，增加英特尔傲腾持久内存的缓存容量可以明显提升查询性能，而增加 DRAM 缓存效果并不明显。在缓存足够多的时候，英特尔傲腾持久内存查询效率和 DRAM 效果接近。而由于英特尔傲腾持久内存成本更低，因此具备更卓越性价比。

第三组测试模拟生产提交重复的 SQL 时，英特尔傲腾持久内存对查询性能的提升效果。本组测试分为 Parquet only、Parquet + 英特尔傲腾持久内存 cache、Parquet + DRAM cache 三种模式，每组测试顺序执行 1 个 IO 较高的 SQL，共执行三次，记录每一种模式每一次的 SQL 执行时间，测试数据如图 4 所示：

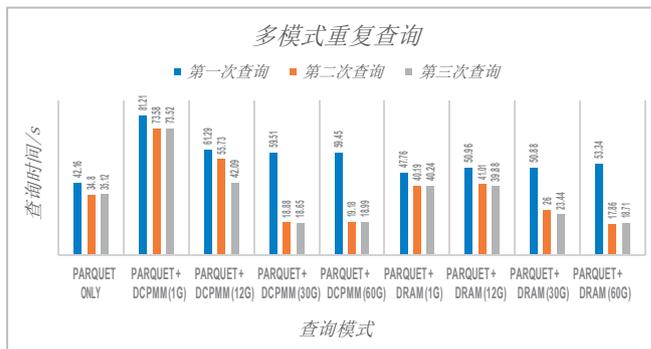


图 4. 多模式重复查询测试 (单位: 秒, 越低越好)

测试结果显示，在重复查询场景较多的环境中，设置缓存能够发挥较好的优化作用。对比最后一次查询时间，英特尔傲腾持久内存 cache 效率提升达到 47%。此外，Parquet 在重复查询 SQL 的场景下能取得优秀的性能提升，英特尔傲腾持久内存和 DRAM 的性能提升效果接近。但是考虑到英特尔傲腾持久内存的价格优势以及扩展容量优势，英特尔傲腾持久内存更适合用于数据缓存层。

第四组测试模拟生产提交多个 SQL 时，英特尔傲腾持久内存缓存对查询性能的提升效果。测试按照 Spark 资源配置的不同，共设置了 8 组测试，其中，测试组 1 和测试组 8 没有使

用英特尔傲腾持久内存缓存，测试组 2 到测试组 8 依次增加英特尔傲腾持久内存缓存，测试人员按执行顺序记录每次 SQL 查询时间总和，测试数据如图 5 所示：

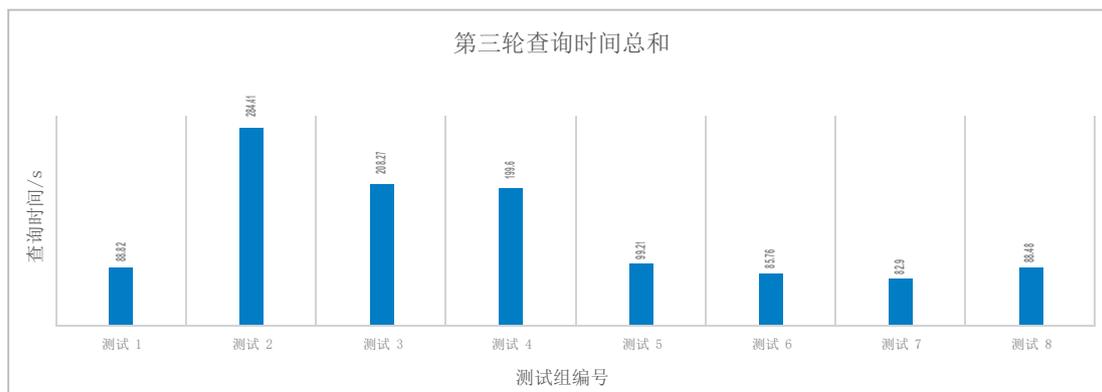


图 5. SQL 查询时间总和测试 (单位: 秒, 越低越好)

测试数据显示，只有使用足够容量的英特尔傲腾持久内存缓存数据，才能提升查询速度，但当缓存容量接近数据量后，继续增加英特尔傲腾持久内存的缓存容量不会对速度有显著提升。

最后一组测试模拟生产提交重复的 SQL 时，英特尔傲腾持久内存缓存对查询性能的提升效果。本轮测试按照 Spark 资源

配置的不同，共设置了 9 组测试，其中，测试组 1 和测试组 9 没有使用英特尔傲腾持久内存缓存，测试组 2 到测试组 8 依次增加英特尔傲腾持久内存缓存数量，并进行 3 轮查询，记录 SQL 执行时间。测试数据如图 6 所示：

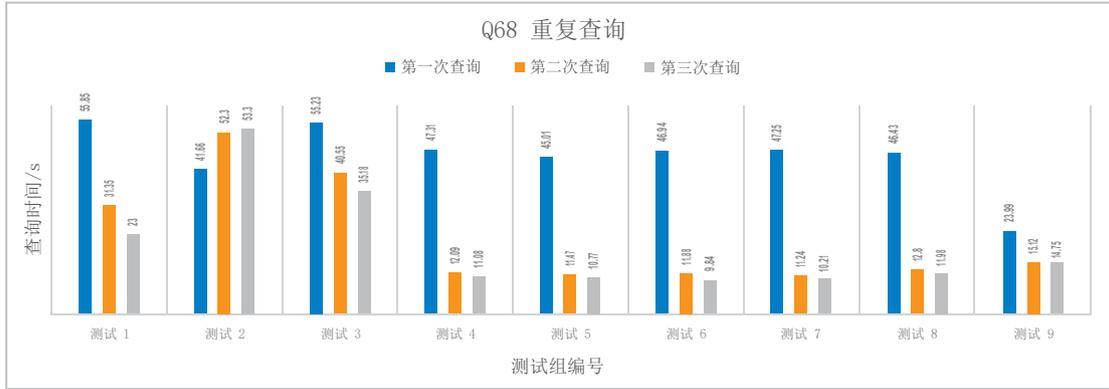


图 6. Q68 重复查询测试（单位：秒，越低越好）

测试数据显示，增加内存和英特尔傲腾持久内存缓存都能提升查询速度。在英特尔傲腾持久内存缓存较小的时候，因为命中率低的问题不会达到理想的查询速度。

从以上测试可以发现，英特尔傲腾持久内存会带来 BEH 性能的较大幅度提升，是因为英特尔傲腾持久内存的性能远高于普通固态硬盘，用作缓存时可以避免数据落在性能较低的固态硬盘上所导致的 IO 瓶颈，降低 CPU 的负载。值得一提的是，在系统性能大幅提升的前提下，硬件的成本却只出现了较小幅度的增长，综合性价比优势十分明显。

### 展望：BEH 加速数据变革，助力企业挖掘大数据价值

东方国信大数据平台企业版 BEH 针对企业应用场景，在高可用、高性能、重要功能特性以及运维等方面进行了创新与增强。在使用英特尔傲腾持久内存优化缓存性能之后，东方国信大数据平台企业版 BEH 的性能得到进一步提升，有望帮助金融、互联网、电信等行业用户有效应对复杂查询及大数据量分析场景，显著改善应用延迟等方面的表现。

“数据处理性能是大数据应用中的重要痛点，英特尔傲腾持久内存展现了在高并发场景下的卓越性能加速效果，同时实现了卓越的成本效益。展望未来，我们将进一步与英特尔等合作伙伴强化合作，完善大数据和云计算技术体系，升级产品结构和产品层次，完善面向大数据采集、汇聚、处理、存储、分析、挖掘、应用、管控为一体的大数据核心能力平台，为行业用户提供更优秀的大数据平台解决方案。”

— 王虎

东方国信技术中心总经理