

代理向量大小超过共享 L2 Cache 时,查询执行时间增速提高,但其增长速度显著低于 NPO 与 PRO.

在最大数据集连接操作中,NPO 算法在 Phi 协处理器上的性能优于 CPU,连接总时间分别为 1 227 785 μ s 和 1 353 713 μ s,PRO 算法在 Phi 和 CPU 上的执行时间为 1 188 729 μ s 和 841 234 μ s,AIR 算法在 Phi 和 CPU 上的执行时间为 407 087 μ s 和 327 758 μ s,PRO 与 AIR 算法在 Phi 上的查询性能略低于 CPU.需要注意的是:实验中使用两块 Intel Xeon E5-2650 v3@2.30GHz 10 核 CPU 和一块 Intel Xeon Phi 5110P@1.053 GHz60 核协处理器,以接近的价格获得了相近的连接操作性能.

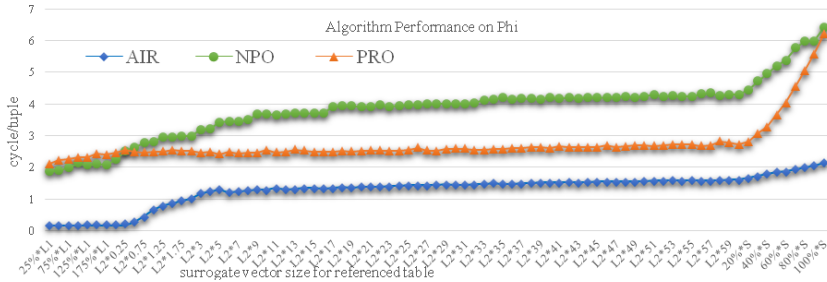


Fig.7 Comparison of different join algorithms on Phi coprocessor

图 7 Phi 协处理器端连接算法性能对比

4 结束语

随着 Xeon Phi 协处理器成为高性能计算的主流处理器,Xeon Phi 协处理器成为连接优化技术研究的新平台.本文的贡献体现在 3 个方面:首先,面向 OLAP 模式特点提出了代理键索引机制,并详细阐述了代理键索引机制的实现及维护技术;然后,通过深入的实验,揭示了基于代理向量参照访问的 AIR 算法的 Cache-Friendly 特征,揭示了 AIR 算法融合了 NPO 小数据和 PRO 大数据时的连接性能特征,验证 AIR 优于 NPO 和 PRO 算法的性能;最后,通过 Phi 协处理器平台上全面的性能对比测试,展现了 3 个算法各自的性能曲线特征,清晰地描述了各个算法在不同连接表大小下的性能特征,为 Phi 协处理器 OLAP 查询处理中连接算法的技术选择提供了详实的实验数据.

从整体实验结果来看,AIR 算法通过 OLAP 模式特点简化连接操作实现技术,通过代理向量参照访问减少内存随机访问数量和 CPU 计算代价,其算法性能与 Cache 大小、并行执行线程数量等硬件参数相关度较高,既能够适应多核 CPU 较大容量 Cache 的优化,也能适应众核协处理器较高并发线程的优化,具有较好的适应性.从总体来说,AIR 算法特点更加适合 Xeon Phi 众核协处理器弱化 Cache 和核心性能、强化并发多线程内存访问的特征,基于代理向量访问的操作简化了算法设计,提高了算法代码执行效率,并且相对于 NPO,PRO 等经典的数据库连接算法更加易于代码改写与平台迁移.

在未来的工作中,我们将继续探索在 Phi 协处理器平台上全面的 OLAP 查询优化技术,尤其是面向 PCI-E 带宽性能瓶颈的 CPU-Phi 协同 OLAP 优化技术.

References:

- [1] <http://www.expreview.com/44000.html>
- [2] Jha S, He BS, Lu M, Cheng XT, Huynh HP. Improving main memory Hash joins on Intel Xeon Phi processors: An experimental approach. Proc. of the VLDB Endowment, 2015,8(6):642–653. [doi: 10.14778/2735703.2735704]
- [3] Polychroniou O, Raghavan A, Ross KA. Rethinking SIMD vectorization for in-memory databases. In: Proc. of the SIGMOD. New York: ACM Press, 2015. 1493–1508. [doi: 10.1145/2723372.2747645]
- [4] Halstead RJ, Absalyamov I, Najjar WA, Tsotras VJ. FPGA-Based multithreading for in-memory Hash joins. In: Proc. of the CIDR. 2015.

- [5] He J, Lu M, He B. Revisiting co-processing for hash joins on the coupled CPU-GPU architecture. Proc. of the VLDB Endowment, 2013,6(10):889–900. [doi: 10.14778/2536206.2536216]
- [6] Morgan TP. Intel mates FPGA with future Xeon server chip. 2014. <http://www.enterprisetech.com/2014/06/18/intel-mates-fpga-future-xeon-server-chip/>
- [7] Blanas S, Li Y, Patel JM. Design and evaluation of main memory Hash join algorithms for multi-core CPUs. In: Proc. of the SIGMOD. New York: ACM Press, 2011. 37–48. [doi: 10.1145/1989323.1989328]
- [8] Balkesen C, Teubner J, Alonso G, Ozsu T. Main-Memory Hash joins on multi-core CPUs: Tuning to the underlying hardware. In: Proc. of the ICDE. Washington: IEEE Computer Society, 2013. 362–373. [doi: 10.1109/ICDE.2013.6544839]
- [9] Albutiu MC, Kemper A, Neumann T. Massively parallel sort-merge joins in main memory multi-core data-base systems. Proc. of the VLDB Endowment, 2012,5(10):1064–1075. [doi: 10.14778/2336664.2336678]
- [10] He BS, Yang K, Fang R, Lu M, Govindaraju NK, Luo Q, Sander P. Relational joins on graphics processors. In: Proc. of the SIGMOD. New York: ACM Press, 2008. 511–524. [doi: 10.1145/1376616.1376670]
- [11] Yuan Y, Lee R, Zhang X. The Yin and Yang of processing data warehousing queries on GPU devices. Proc. of the VLDB Endowment, 2013,6(10):817–828. [doi: 10.14778/2536206.2536210]
- [12] Pirk H, Manegold S, Kersten M. Accelerating foreign-key joins using asymmetric memory channels. In: Bordawekar R, Lang CA, eds. Proc. of the Int'l Workshop on Accelerating Data Management Systems Using Modern Processor and Storage Architectures (ADMS). 2011. 27–35.
- [13] Zhang Y, Zhou X, Zhang Y, Zhang Y, Su M, Wang S. Virtual denormalization via array index reference for main memory OLAP. IEEE Trans. on Knowledge & Data Engineering, 2016,28(4):1061–1074. [doi: 10.1109/TKDE.2015.2499199]
- [14] https://en.wikipedia.org/wiki/Surrogate_key



张宇(1977—),女,黑龙江绥化人,博士,副教授,主要研究领域为 GPU 数据库,数据仓库,OLAP.



陈红(1965—),女,博士,教授,博士生导师,CCF 高级会员,主要研究领域为数据仓库,数据挖掘,传感器网络.



张延松(1973—),男,博士,副教授,主要研究领域为内存数据库,数据仓库,OLAP.



王珊(1944—),女,教授,博士生导师,CCF 会士,主要研究领域为高性能数据库,内存数据库,数据仓库.