arXiv:2601.12123v1 [cs.DB] 17 Jan 2026

量子计算是否已准备好进行实时数据库优化?

刘瀚文计算机科学系 南加州大学洛杉矶，加利福尼亚州，美国
hanwen liu@usc.edu

Ibrahim Sabek计算机科学系 南加州大学洛杉矶，加利福尼亚州，美国 sabek@usc.edu

–

I. 背景/问题/方法。

数据库系统包含多个对性能至关重要的优化任务，例如连接排序（如 [1]–[8]）和索引调优（如 [9]–[11]）。随着数据量不断增长、工作负载日益复杂，这些问题要实现高效求解变得愈发困难，且难度呈指数级增长 [12]。量子计算，尤其是量子退火 [13], , 是一种很有前景的范式，能够通过量子隧穿 [16] 高效探索超大规模搜索空间 [14], [15] 。如图 1 所示，它无需跨越能垒，而是通过隧穿穿过能垒，从而摆脱局部最优。早期工作 [17]–[24] 主要致力于为数据库优化问题（例如连接顺序 [1]）提供抽象表示（如二次无约束二进制优化（QUBO） [25]），但由于量子计算服务的高昂开销（例如 D-Wave 的 CQM-Solver [26]最低运行时间为 5 秒），忽略了与数据库系统的实际集成。近期，量子退火服务提供商推出了延迟更低的解决方案，例如 NL-Solver [27], , 为在数据库管理系统中真正实现量子解决方案铺平了道路 [28]。然而，这也带来了新的系统研究挑战，即如何平衡效率与解决方案质量。

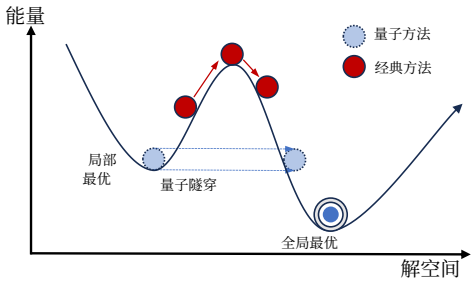


图 1：量子退火与经典方法的比较。

在本次演讲中，我们将展示如何实现这种平衡。作为概念验证，我们介绍Q^2O，即首个实时量子增强型查询优化器。图 2 展示了端到端工作流程：我们利用实际的数据库统计信息，将连接顺序问题编码为非线性模型，这种格式可由 NL-Solver 求解；随后将求解结果转换为计划提示

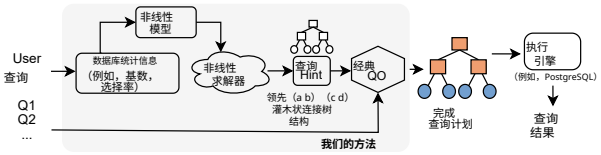


图 2：我们提出的 Q^2O 框架的工作流程。

以指导 PostgreSQL 的查询优化器生成完整计划。Q^2O 能够实时处理实际查询。

II. 结果/结论

我们搭建了 PostgreSQL 16.4，并配套使用*pg hint plan* 版本。我们使用Q^2O，以连接顺序基准测试（*JOB*） [29] 作为现实世界的优化挑战进行评估。**查询计划质量**。我们评估生成的

通过测量 PostgreSQL 上 JOB 工作负载的执行时间来评估查询计划。图 3 将我们的方法生成的计划与 PostgreSQL 的默认查询计划进行了比较。在 113 个查询中，我们的方法使 31 个查询获得了加速；这些查询的延迟最高降低了 92.7%，平均降低了 42.09%。

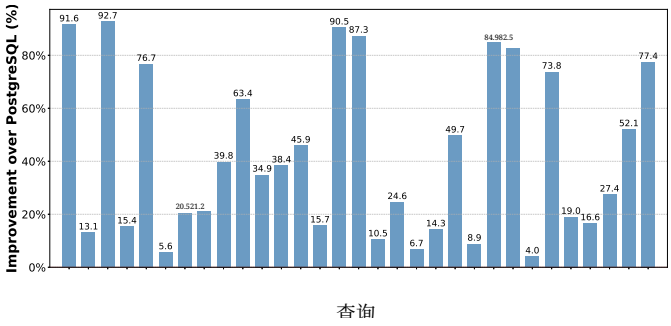


图 3：我们的方法生成的部分查询计划优于 PostgreSQL 的默认计划。

端到端（E2E）延迟。我们测量整个流水线的端到端延迟。对于 PostgreSQL，该延迟包括规划时间和执行时间；对于 Q^2O，则包括 PostgreSQL 规划（含提示）、PostgreSQL 执行以及 NL-Solver 耗时。表 I 报告了四条 JOB 查询各组件的延迟。

表 I：端到端延迟细分：我们的方法与 PostgreSQL 的对比。

	时间（毫秒）	q21	q60	q62	q63
PG	规划	1.00	3.14	3.17	1.16
	执行	3581.40	10951.31	4680.26	4846.42
Ours	带提示的规划	2.24	8.31	9.03	9.02
	执行	272.35	6010.00	429.58	585.60
	NL-Solver	2530.22	2748.32	2854.31	2816.42
Gain	**执行时间 ↑**	**13.15x**	**1.82x**	**10.89x**	**8.28x**
	端到端延迟 ↑	**1.28x**	**1.25x**	**1.42x**	**1.42 倍**

尽管 NL-Solver 会带来额外开销（主要是无法避免的云通信），我们的方法仍能生成质量更高的连接顺序方案，使执行时间最高提速 13.15×，并降低整体端到端延迟。

总之，我们提出了Q^2O，利用量子计算解决实时场景下的数据库优化问题。我们的评估表明，查询执行时间和端到端延迟均得到了加速。展望未来，该框架有望应对实时场景中规模更大、更加复杂的数据库优化问题。

参考文献

[1] Y. E. Ioannidis, 《查询优化》, *ACM Comput. Surv.*, 第28卷, 第1期, 第121–123页, 1996年3月。 [在线]。获取地址: https://doi.org/10.1145/234313.234367[2] J. Zhao、K. Su、Y. Yang、X. Yu、P. Koutris 和 H. Zhang, 《揭穿连接排序的神话：迈向稳健的SQL分析》, *Proc. ACM Manag. Data*, 第3卷, 第3期, 2025年6月。 [在线]。获取地址: https://doi.org/10 1145 3725283 . / R. Marcus、P. Negi、H. Mao、N. Tatbul、[3] M. Alizadeh 和 T. Kraska, 《Bao：让学习式查询优化切实可行》, 载于 *SIGMOD*, 2021年。[4] X. Yu、G. Li、C. Chai 和 N. Tang, 《利用树形LSTM进行强化学习，以选择连接顺序》, 载于*2020*年*IEEE*第*36*届国际数据工程会议*(ICDE)*。IEEE, 2020年, 第1297–1308页。[5] R. Marcus 和 O. Papaemmanouil, 《用于连接顺序枚举的深度强化学习》, 载于第一届利用人工智能技术进行数据管理国际研讨会论文集, 丛书 aiDM'18。美国纽约州纽约市: 计算机协会, 2018年。 [在线]。获取地址: https://doi.org/10.1145/3211954.3211957[6] H. Liu、S. Giridhara 和 I. Sabek, 《用于可验证学习式查询优化的保形预测》, 载于超大型数据库国际会议论文集, *VLDB*, 英国伦敦, 2025年。[7] H. Liu、Q. Zhang、R. Marcus 和 I. Sabek, 《Serag：用于查询优化的自演进RAG系统》, 2025年。 [在线]。获取地址: https://viterbi-web.usc.edu/~sabek/pdf/25 workshop serag.pdf [8] ——, 《Sefrqo：一种自演进的微调基于RAG的查询优化器》, 2025 [ ] 2508 17556 。在线。获取地址: [9] https://arxiv.org/abs/ 。S. Agrawal、S. Chaudhuri 和 V. R. Narasayya, 《SQL数据库中物化视图和索引的自动选择》, 载于第*26*届超大型数据库国际会议论文集, 丛书 VLDB '00。美国加利福尼亚州旧金山: Morgan Kaufmann 出版公司, 2000年, 第496–505页。[10] S. Chaudhuri 和 V. R. Narasayya, 《适用于Microsoft SQL Server的高效成本驱动索引选择工具》, 载于第*23*届超大型数据库国际会议论文集, 丛书 VLDB '97。美国加利福尼亚州旧金山: Morgan Kaufmann 出版公司, 1997年, 第146–155页。[11] W. Wu、C. Wang、T. Siddiqui、J. Wang、V. Narasayya、S. Chaudhuri 和 P. A. Bernstein, 《利用强化学习进行预算感知索引调优》, 载于*2022*年国际数据管理会议论文集, 丛书 SIGMOD '22。美国纽约州纽约市: 计算机协会, 2022年, 第1528–1541页。 [在线]。获取地址: https://doi.org/ 10 1145 3514221 3526128 . / 。D. [12] Shasha 和 P. Bonnet, 《数据库调优：原理、实验与故障排除技术》。美国加利福尼亚州旧金山: Morgan Kaufmann 出版公司, 2002年。[13] T. Kadowaki 和 H. Nishimori, 《横场伊辛模型中的量子退火》, Phys. Rev. E, 第58卷, 第5355–5363页, 1998年11月。 [ ] 10 1103 58 5355 在线。获 [14] 取地址: https://link.aps.org/doi/ 。/PhysRevE. 。A. Awasthi、N. Kraus 等, 《量子—经典优化在生产调度中的实际应用》, 载于2024年IEEE国际量子计算与工程会议（QCE）, 第02卷, 2024年, 第239–244页。[15] G. Colucci 等, 《电力网络优化：一种量子方法》, IEEE Access, 第11卷, 第98 926–98 938页, 2023年。[16] M. Razavy, 《隧穿量子理论》。世界科技出版公司, 2013年。[17] P. Saxena、I. Sabek 和 F. Spedalieri, 《用于优化连接顺序的约束二次模型》, 载于第1届量子计算与量子启发技术在数据密集型系统及应用中的应用研讨会论文集, 丛书 Q-Data ' 24, 2024, 38 44 [ ] 2 p. – 。在线。获取地址: https://doi-org.libproxy .usc.edu/ 10.1145/3665225.3665447[18] H. Liu、P. Saxena、F. Spedalieri 和 I. Sabek, 《通过约束二次模型优化连接顺序（摘要）》, 载于第1届量子数据与机器学习研讨会论文集, 丛书 QDML ' 25, 2025年。[19] H. Liu、F. Spedalieri 和 I. Sabek, 《Q2O演示：量子增强查询优化器》, 载于超大型数据库国际会议论文集, VLDB, 2025年。[20] S. Groppe 和 J. Groppe, 《通过为Grover搜索算法生成代码，在通用量子计算机上优化事务调度》, 载于第25届国际数据库工程与应用研讨会论文集, 丛书 IDEAS '21。美国纽约州纽约市: 计算机协会, 2021年, 第149–156页。 [ ] 10 1145 3472163 3472164 在线。获取地址: [21] https://doi.org/ 。/ 。N. Nayak、A. Prisacaru、U. C.alıkyılmaz、J. Groppe 和 S. Groppe, 《通过使用锁定机制迭代求解QUBO，以降低复杂度实现量子增强事务调度》, 载于第2届量子计算与量子启发技术在数据密集型系统及应用中的应用研讨会论文集, 丛书 Q-Data '25。美国纽约州纽约市: 计算机协会, 2025年, 第 26 35 [ ] 10 1145 3736393 3736701 – 。在线。 [22] 获取地址: https://doi.org/ 。/ 。I. Trummer 和 C. Koch, 《在D-Wave 2X绝热量子计算机上进行多查询优化》, 载于VLDB, 2016年。[23] T. Fankhauser、M. E. Sol`er、R. M. F¨uchslin 和 K. Stockinger, 《利用经典与量子计算混合方法进行多查询优化》, arXiv预印本 arXiv:2107.10508, 2021年。[24] M. Sch¨onberger、I. Trummer 和 W. Mauerer, 《用于连接排序的量子启发数字退火》, 载于VLDB, 2023年。[25] A. Lucas, 《许多NP问题的伊辛形式》, Frontiers in Physics, 第2卷, 2014年。 [在线]。获取地址: https://www.frontiersin.org/ articles/10.3389/fphy.2014.00005[26] 约束二次模型混合求解器 [白皮书], https://www.dwavesys.com/media/rldh2ghw/14-1055a-a hybrid solver for constrained quadratic models.pdf, 2021年。[27] E. Osaba 和 P. Miranda-Rodriguez, 《D-Wave非线性规划混合求解器：描述与性能分析》, IEEE Access, 第13卷, 第4724–4736页, 2025年。[28] H. Liu、A. Kumar、F. Spedalieri 和 I. Sabek, 《用于灌木状连接树的量子—经典混合优化》, 载于第2届量子计算与量子启发技术在数据密集型系统及应用中的应用研讨会论文集, 丛书 Q-Data '25。美国纽约州纽约市: 计算机协会, 2025年, 第 20 24 [ ] 10 1145 3736393 3736695 – 。在线。 [29] 获取地址: https://doi.org/ 。/ 。V. Leis、A. Gubichev、A. Mirchev、P. Boncz、A. Kemper 和 T. Neumann, 《查询优化器究竟有多出色？》, Proc. VLDB Endow., 第9卷, 第3期, 第204–215页, 2015年11月。 [在线]。获取地址: https://doi.org/10.14778/2850583.2850594