迈向 Hyperledger Fabric 的最高效状态数据库

Ivan Laishevskiy Artem Barger Vladimir Gorgadze*Idea Blockchain Research Lab Idea Blockchain Research Lab* 莫斯科物理技术学院莫斯科，俄罗斯 海法，以色列 莫斯科，俄罗斯 ivan.laishevskii@scientificideas.org bartem@scientificideas.org gorgadze.vv@mipt.ru

arXiv:2303.08875v1 [cs.DC] 15 Mar 2023

摘要——**Hyperledger Fabric 是广为人知且最具代表性的企业级许可型区块链。Hyperledger Fabric 的架构引入了模拟-排序-验证这一全新架构范式以及可插拔架构，从而允许更高程度的定制。其中一个关键组件是世界状态数据库，负责捕获区块链应用状态的快照。Hyperledger Fabric 通过键值数据库抽象来管理状态，对等节点会在交易经过验证后更新状态，并在模拟过程中读取状态。因此，提供良好的读写性能会影响系统的整体性能。目前，Hyperledger Fabric 支持两种不同的状态数据库实现：一种是基于 LSM树的嵌入式数据库 LevelDB，另一种是 CouchDB。本研究着重于探寻状态数据库的替代实现，并分析是否存在更优且更具扩展性的方案。我们评估了多种可接入 Hyperledger Fabric 的数据库，如 RocksDB、BoltDB 和 BadgerDB。本文介绍了如何接入新的状态数据库，并基于各种工作负载给出性能测试结果。**

索引词——**区块链、键值存储、数据库、Lev- elDB、RocksDB、BadgerDB、BoltDB、Hyperledger Fabric、Hy- perledger Caliper、状态数据库。**

I. 引言

区块链技术日益普及 [1]–[4]，为企业开始采用基于区块链的解决方案铺平了道路，尤其是采用 Hyper- ledger Fabric [5]–[7] 实现的解决方案。Hyperledger Fabric（HLF）是一项开源计划，致力于构建企业级许可型区块链平台 [8], [9]。

状态数据库是 HLF 最重要的组成部分之一（StateDB）。StateDB 存储区块链网络状态的最新快照。HLF 对等节点依靠 StateDB 读取当前值及其版本，以构建交易的读集。随后，在提交与验证阶段执行多版本并发控制 [11] 检查时，对等节点会从 StateDB 读取数据，以比较读集版本。最后，验证完成后，对等节点将已验证的交易提交到 StateDB，以反映最新变更。 [10]

根据 [12], , HLF 中的提交阶段占据了交易处理时间的大部分，因此成为性能瓶颈。此外，该研究表明，对等节点复制交易时，读取操作发挥着重要作用。显然，与 StateDB 的交互会直接影响 HLF 的整体性能。目前，HLF 提供了两种 StateDB 实现方案，一种基于 GoLevelDB，另一种基于 CouchDB。因此，Fabric社区认识到有必要开发一种更优的 StateDB 替代方案 [13]。本文将分析 StateDB 的各种选项，并考察基于 LSM 树或 B+树 [14], [15]的嵌入式数据库的各种实现。

II. 背景

我们需要建立一个比较基准，因此实现了一项基准测试，以确定用性能更高的替代方案取代当前实现所能带来的最大潜在性能提升。该基准测试实现了链码 FixedAsset，其中包含两个方法：一个会使对等节点与 StateDB 交互，另一个则不会。

由于已有结果表明 StateDB goleveldb 的性能优于 CouchDB [16]，我们使用 StateDB goleveldb 来考察 StateDB 对 HLF 性能的总体影响。为此，我们运行了多项性能基准测试。

根据图 1 中平均每秒事务数的比较结果，我们可以得出结论：与向 StateDB 写入相同的值相比，在交易中从 StateDB 读取 100 字节的值，对 HLF 性能几乎没有影响。

因此，本研究将提升 StateDB 的写入性能作为首要目标。

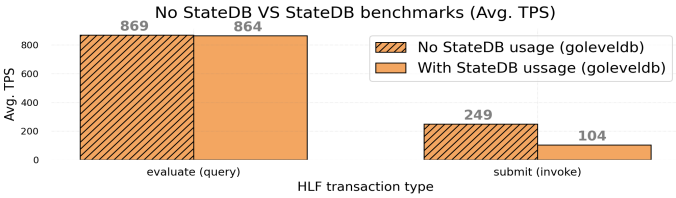


图 1. 无 StateDB 与使用 StateDB 的性能基准测试结果。

III. 集成替代状态数据库

在开源 HLF 代码（提交 [17]）中，已确定负责 HLF 与 StateDB 交互的主要组件

979-8-3503-1019-1/23/$31.00 ©2023 IEEE

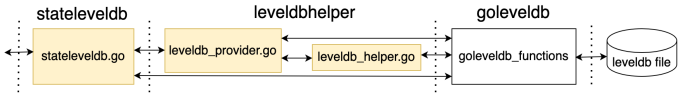


图2. HLF中负责与StateDB交互的组件（以颜色突出显示）。

。这些组件如图2所示。为了添加新的状态数据库，我们决定为每种数据库创建与标记组件类似的组件。因此，主要的代码升级是为每种数据库添加三个代码文件：stateNameOfDB.go、NameOfDB provider.go、NameOfDB helper.go。goleveldb 的功能得到了完整保留。要选择 StateDB，需在运行容器前于 core.yaml 文件中指定所选数据库的名称。

除了所述模块外，我们还创建了覆盖新增代码的测试文件，并对其他一些HLF模块进行了一些小幅修改。

IV. HLF中嵌入式StateDB基准测试

A. 性能基准测试评估配置

我们使用Hyperledger Caliper工具[21], 对采用新嵌入式状态数据库的HLF进行了性能评估。该工具为多个区块链平台（包括Hyperledger Fabric）提供通用基准测试。

该工具支持采集基本的区块链性能 [22] 指标。由于需要比较的指标数量庞大，为了简化分析并便于可视化，本文仅考察写事务吞吐量。

性能基准测试基于来自 [23] 的代码示例。

运行性能基准测试的对等节点机器配置如下：操作系统：Ubuntu 20.04.4 LTS，CPU：4，内存：16，总固态硬盘容量：60 GB，最大带宽（读 — 写）：30 MB/s — 30 MB/s，最大 IOPS（读 — 写）：2000 — 2000，CPU 系列：6，型号：106，CPU 名称：Intel Xeon处理器（Icelake），每核线程数：2，每个插槽的核心数：2，插槽数：1。

FixedAsset链码方法 [24] 使用 Golang 开发，用于评估与 StateDB 进行各种交互的性能。

*B. Create-Asset*性能基准测试

本研究的场景旨在向 StateDB 写入数据。写入性能基准测试包括对 FixedAsset 链码的 CreateAsset方法进行 Submit 调用。我们将此类调用称为交易。该链码分别部署在各个独立的 HLF 网络中，每个网络使用一种待考察的 StateDB：RocksDB、bboltdb 或 BadgerDB。

每笔交易向 StateDB 写入一个键值对。对于每种交易类型，每个客户端运行 5 分钟：100 字节 /1000 字节 / 4000 字节 / 8000 字节 / 16000 字节 / 24000 字节 / 32000 字节 / 64000 字节。

图3比较了嵌入式状态数据库处理 HLF 交易的吞吐量。结果显示，所有状态数据库之间的竞争几乎势均力敌。不过，BadgerDB 状态数据库在各类交易中的每秒事务数略占优势。此外，对于 64KB 的交易大小，BadgerDB 的每秒事务数远高于 goleveldb。另一方面，RocksDB 的性能与 goleveldb 持平，并且在交易大小为 64KB 时表现出明显优势。类似地，对于大小为 32KB 和 64KB 的交易，bbolt 相较于 goleveldb 仅表现出微弱优势。

因此，我们可以说，BadgerDB 的写入每秒事务数高于其他状态数据库。

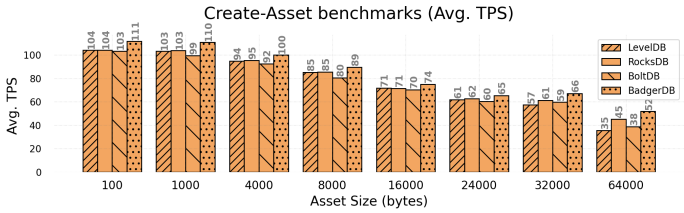


图3. 创建资产性能基准测试结果（每秒交易数）。

五、结论与后续研究方向

研究选取了几个在性能和功能方面可与 goleveldb 竞争的数据库，并将其作为 StateDB 集成到 HLF 中。此外，还为每种可能的替代 StateDB 创建并执行了性能基准测试。最终，BadgerDB 被确定为最具潜力的 StateDB。

BadgerDB 在 StateDB 写入性能基准测试中展现出相当明显的优势。在写入 64KB 数据时，BadgerDB 相较于 goleveldb 的优势最为突出（每秒事务数几乎高出 1.5 倍）。此外，BadgerDB 还为 StateDB 提供了 goleveldb 尚未实现的功能：它保证 ACID 属性，并可借助额外工具 [25]在 HLF 中使用自定义查询。

包含所链接数据库的 HLF 源代码可在以下代码仓库中获取：[18] (RocksDB)、 [19](bbolt)、 [20] (Bad- gerDB)。

为运行性能基准测试，我们修改了 HLF fabric-samples/test-network，以提供测试新数据库的 StateDB 对等节点性能所需的最小网络组件集：https: // github.com/fubss/fabric-samples/tree/dbs selector。

基于 Hyperledger Caliper [23] 演示，我们为 HLF 负载创建了一个自定义代码仓库，并针对 Linux 和 OS X 进行了适配（位于不同分支中）：https://github.com/ fubss/caliper-workspace-3。

后续研究建议：

- 比较嵌入式数据库的其他性能指标；
- 开展读工作负载性能基准测试；
- 找出并消除其他拖慢它的 HLF组件。

致谢

我们感谢 Vladimir Chechetkin 为本文的工作集成 BadgerDB 并作出其他贡献。

参考文献

[1] Y. Manevich、A. Barger 和 G. Assa，《从执行-排序-验证区块链中删除交易》，载于 2021年 IEEE 国际区块链与加密货币会议（ICBC）。IEEE，2021年，第 1–9 页。[2] J. Abou Jaoude 和 R. George Saade，《区块链应用——在不同领域的应用》，IEEE Access，第 7 卷，第 45 360–45 381 页，2019年。[3] S. M. Nor、M. Abdul-Majid 和 S. N. Esrati，《区块链技术在促进伊斯兰社会金融方面的作用：马来西亚天课管理案例》，foresight，2021年。[4] I. Fedorov、A. Pimenov、G. Panin 和 S. Bezzateev，《5G网络中的区块链：私有区块链的性能评估》，载于 2021年波电子学及其在信息与电信系统中的应用会议（WECONF）。IEEE，2021年，第 1–4 页。[5] Hyperledger基金会，《面向企业的区块链平台》，2022年，修订版 dd63f081。 [在线]。网址：https://hyperledger-fabric. readthedocs.io/en/latest/[6] ——。（2022年）案例研究。 [在线]。网址：https://www.hyperledger. org/learn/case-studies [7] A. Barger、O. Ilina、A. Zemtsov 和 K. Tagirova，《基于 Hyperledger Fabric 的可信慈善基金会平台》，载于 2022年 IEEE 国际全层智能系统会议（COINS）。IEEE，2022年，第 1–6 页。[8] Hyperledger基金会。（2022年）Hyperledger Fabric。 [在线]。网址：https://www.hyperledger.org/use/fabric [9] E. Androulaki、A. Barger、V. Bortnikov、C. Cachin、K. Christidis、A. De Caro、D. Enyeart、C. Ferris、G. Laventman、Y. Manevich 等，《Hyperledger Fabric：面向许可型区块链的分布式操作系统》，载于第十三届 EuroSys会议论文集，2018年，第 1–15 页。[10] A. Barger、Y. Manevich、H. Meir 和 Y. Tock，《用于 Hyperledger Fabric 的拜占庭容错共识库》，载于 2021年 IEEE 国际区块链与加密货币会议（ICBC）。IEEE，2021年，第 1–9 页。[11] C. H. Papadimitriou 和 P. C. Kanellakis，《关于基于多版本的并发控制》，ACM数据库系统汇刊（TODS），第 9 卷，第 1 期，第 89–99 页，1984年。[12] T. Nakaike、Q. Zhang、Y. Ueda、T. Inagaki 和 M. Ohara，《使用 goleveldb基准测试对 Hyperledger Fabric 进行性能特征分析与优化》，载于 2020年 IEEE 国际区块链与加密货币会议（ICBC）。IEEE，2020年，第 1–9 页。[13] （2022年）Go 语言中的 LevelDB 键值数据库。GitHub 仓库。 [在线]。网址：https://github.com/syndtr/goleveldb [14] D. Enyeart。（2021年）Fabric战略优先事项。 [在线]。网址：https://lists.hyperledger.org/g/fabric/message/10357/[15] 2021 Hyperledger基金会。（）Fabric战略优先事项——2021年讨论。 [在线]。网址：https://wiki.hyperledger.org/display/ fabric/Fabric+Strategic+Priorities+-+2021+discussion [16] P. Thakkar、S. Nathan 和 B. Viswanathan，《Hyperledger Fabric区块链平台的性能基准测试与优化》，载于 2018年 IEEE 第 26 届计算机与电信系统建模、分析和仿真国际研讨会（MASCOTS）。IEEE，2018年，第 264–276 页。[17] （2021年）节点链码的 Evaluate()错误响应。GitHub 仓库中的提交。 [在线]。网址：https://github.com/hyperledger/ fabric/commit/6656f72563c73a806dee7068dd91b3acf2a286aa[18] （2022年）采用嵌入式 grocksdb 的 Hyperledger Fabric。GitHub 仓库中的分支。 [在线]。网址：https://github.com/fubss/fabric/ tree/grocksdb-[19] 2022 （）采用嵌入式 bbolt 的 Hyperledger Fabric。GitHub 仓库中的分支。 [ ] 30 在线。网[20] 2022 址：https://github.com/fubss/fabric/tree/bbolt- （）采用嵌入式 BadgerDB 的 Hyperledger Fabric。GitHub 仓库中的分支。 [在线]。网址：https://github.com/fubss/fabric/ tree/badger-[21] 2022 （）Hyperledger Caliper。GitHub 仓库。 [在线]。网址：https://github.com/hyperledger/caliper [22] 性能与规模工作组（PSWG）。（2018年）关键指标定义。文章。 [在线]。网址：https://www.hyperledger.org/learn/publications/blockchain-performance-metrics#definitions [23] （2022年）Hyperledger Caliper 基准测试。GitHub 仓库。 [在线]。网址：https://github.com/hyperledger/caliper-benchmarks [24] （2022年）FixedAssetContract。GitHub 仓库中的文件。[在线]。网址：https://github.com/fubss/caliper-workspace-3/blob/main/smart-contract/go/FixedAssetContract.go [25] （2022年）Badgerhold。GitHub 仓库。 [在线]。网址：https: //github.com/timshannon/badgerhold