Poseidon Chain

Technical WhitePaper

产业级区块链技术平台技术白皮书

目录

区块链发展历程和 POSE 的技术革新

从 2008 年 Bitcoin 项目发布至今，区块链技术发展史上已经出现了四个里程碑式的项目，并普遍认为已经发展到了第二代。POSE 将作为第五个里程碑式的项目，带领区块链迈入第三代。

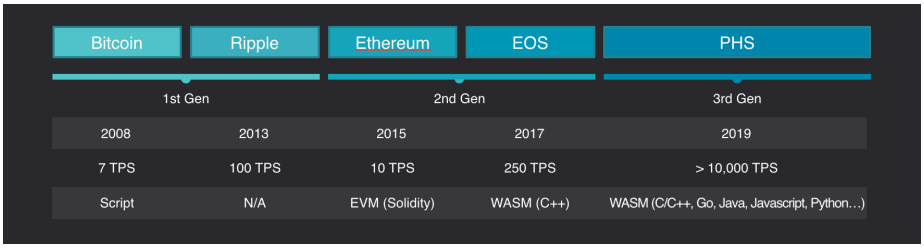

图 1:

在第一代区块链项目里，以 Bitcoin 和 Ripple 为代表，它们完成了第一代的去中心化区块链架构，并将 TPS 逐步优化到了 100 的量级。

在第二代区块链项目里，以 Ethereum 和 EOS 为代表，它们实现了图灵完备的智能合约引擎，为去中心化的区块链技术添加了可编程的属性，并将 TPS 优化到了 250 的量级。

业界普遍认为第三代区块链将迈向可编程社会，POSE 将成为第三代区块链的里程碑项目，实现产业级的区块链平台，TPS 将达到万级以上。

但现代企业的信息化需求是多种多样的，非常复杂。我们在对现代商业众多技术需求分析总结后认为有三点是最主要的需求，他们分别是：

- 可延展性，能够支持广泛的实体产业智能上链，不论体量大小，技术上需要：
  - 海量数据的处理能力
  - 产业级的吞吐性能
- 确定性，或强一致性，让产业在链上低成本运行，包括：
  - 高效治理机制
  - 快速确认且杜绝分叉
  - 动态费率及成本隔离
- 可交互性，帮助更多的开发者和企业更便捷的挖掘链上红利，包括：
  - 开发者友好的 API 和 SDK
  - 多种流行开发语言的支持

POSE 在这三方面都将有着巨大的革新，真正带领区块链迈入第三代，下面我们将逐一说明。

可延展性，容纳广泛的实体产业

POSE 的可延展性通过支持海量数据处理和优异的吞吐性能实现，从而达到目前互联网主流的技术能力水平。

海量数据处理

要支持公证、溯源甚至更复杂的多样的企业商业逻辑，我们必然需要一个强大的通用数据存储和高效的查询技术，而这在区块链上一直是一个非常大的短板。

在处理数据的历史上，基于关系代数的关系型数据处理已经非常成熟，并被广泛的证明了实用价值，不管是数据库和数据仓库中都有着很广泛的应用，这一成熟的数据处理理论体系无疑是最好的选择。

但在区块链上构建关系型数据存储一直有着很大的挑战，历史上曾有众多区块链项目试图提供这一能力，但都并没有取得成功，我们分析其主要原因在于这些项目都在试图引入对标关系型数据库的存储能力，而关系型数据库对数据的整体性和逻辑严密性有着非常高的要求，这和区块链大量离散数据的存储形式存在着本质的冲突。

因此，POSE 将实现一个新的数据处理架构，这个数据模型更适合于区块链上大量的离散数据，但又能很好的提供关系型数据操作，从而从本质上化解这个矛盾。这个架构分为存储和查询两部分实现。

去中心化的关系型存储

在这个新的架构里，POSE 创新性的引入了一个新的通用数据模型。这个模型分为三部分，Schema、Relation 和 Data。Schema 用于描述数据结构，Relation 用于描述数据关系，Data 则为实际的数据存储。这个数据模型本身已经是一个成熟技术，在非结构化数据以及数据仓库里有着广泛的应用并取得了很好的效果，基于这个数据模型提供的关系型查询技术也非常成熟和高效。我们在做了深入的分析和研究后认为，这个模型在区块链上也同样可以取得很好的效果。

下面将详细描述这三部分，这三部分都采用 JSON 格式描述。

1. Schema

Schema 用来描述数据格式，遵循 JSON Schema 规范[1]，能够很好的被人类和程序阅读，有着完整的结构验证，有利于自动化测试，也有利于程序执行，可以方便的验证提交数据的有效性。下面是一个例子。

```
{
  "$id": "a-product-order",
  "$schema": "http://json-schema.org/draft-07/schema#",
  "title": "Order",
  "type": "object",
  "properties": {
    "orderID": {
      "type": "string",
      "description": "The ID of an order."
    },
    "sku": {
```

[1] https://json-schema.org/

```
      "type": "string",
      "description": "The ID of product in this order."
    },
    "amount": {
      "description": "Amount in this order, which must be equal to or greater than zer
      "type": "integer",
      "minimum": 0
    }
  }
}
```

2. Relation

Relation 用来描述数据之间的关联关系，包括主键和外键的定义。主要用于建立索引以及快速检索数据，为 POSE 提供接近关系型数据库的查询能力。

```
{
  "$id": "a-product-order",
  "type": "relation",
  "primaryKey": ["orderID"],
  "foreignKeys": [ {
    "fields": "sku",
    "reference": {
      "id": "a-product-list",
      "fields": "skuID"
    }
  } ]
}
```

3. Data

Data 即实际数据，遵循 JSON 格式。通过应用不同的 Schema 和 Relation，POSE 可以实现完整的数据关系，从而支持复杂的商业逻辑。

高速关系型查询引擎

在这个通用的数据模型基础上，POSE 还将实现一套基于下图所示关系代数的查询引擎，将查询指令通过解释器转换为对实际数据的检索作业在查询引擎中执行，通过优化索引创建，优化数据查询，将元数据在内存中检索，

大大减少了在查询过程中执行语义检查的时间。

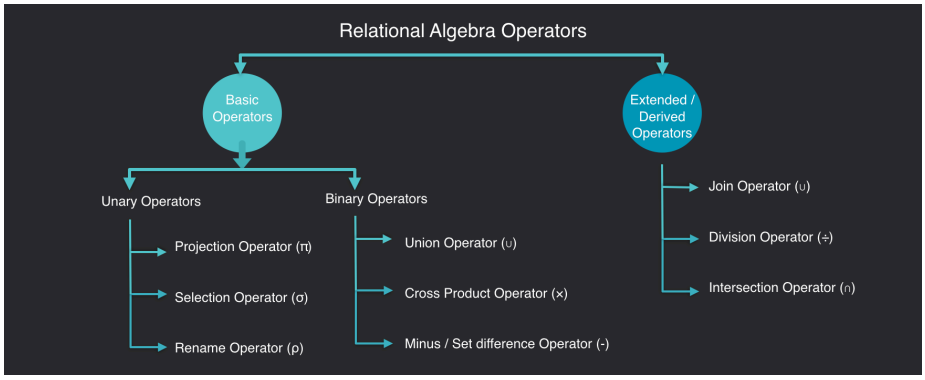

图 2:

这个查询引擎主要分为三个部分

1．用户接口

用户接口接收具体的查询请求，并将其转交给解释器处理。

2．元数据存储

元数据需要在内存中检索，因此我们需要一个高效的内存元数据缓存，元数据包括数据模型的名字，字段间的关联关系，对应的实际数据链上位置等。

3．解释器、执行器

解释器完成查询指令的从词法分析、语法分析、编译、优化以及查询计划的生成。生成的查询计划转交给执行器执行。

结合一个通用的数据模型和一个高效的查询引擎，POSE 可以提供接近关系型数据库的数据处理能力。

产业级的吞吐性能

性能一直是区块链长久以来的瓶颈。目前业界普遍认为，通过数据分片、并行链和 DPoS 等技术，可以降低延迟和提升数据吞吐量。大部分区块链项目都是这么做，但实际效果并不尽如人意。

从数据上看，Bitcoin 是大概最少 30 分钟后数据才能被确认，TPS 约为 7；Ripple 可以在分钟级别确认交易，但 TPS 只有 100 左右；Ethereum 也需要几分钟才能确认数据，由于引入了智能合约，TPS 只有十几；EOS 做了很

多改进，可以在分钟级别确认交易，但 TPS 仍然只有 250[2]，即使是测试网，最好成绩只曾达到 650 的 TPS[3]，离其宣称的百万 TPS 相差甚远。

这些区块链项目里通常遇到的最主要的性能瓶颈是无法进行并发计算的困境，其中，以 Ripple 为例，虽然支持多线程，但只是不同的功能模块之间进行了线程拆分，实际的区块打包仍然是在一个线程中以顺序的方式同步执行。一方面，这样在区块链的核心逻辑区块打包这个任务上并没有提升性能；另一方面，Ripple 引入了大量的线程同步工作，并没有完全发挥出并行计算的优势。

实际上，基本所有的区块链项目都存在着这样的问题。而 EOS 甚至并未计划实现多线程并行[4]，这意味着只能依靠 CPU 单核计算能力的提升来提升性能。如下图所示，在如今多核已经成为主流、高性能计算已经迈入百核量级的现代计算体系中，这个性能瓶颈将愈发凸显。

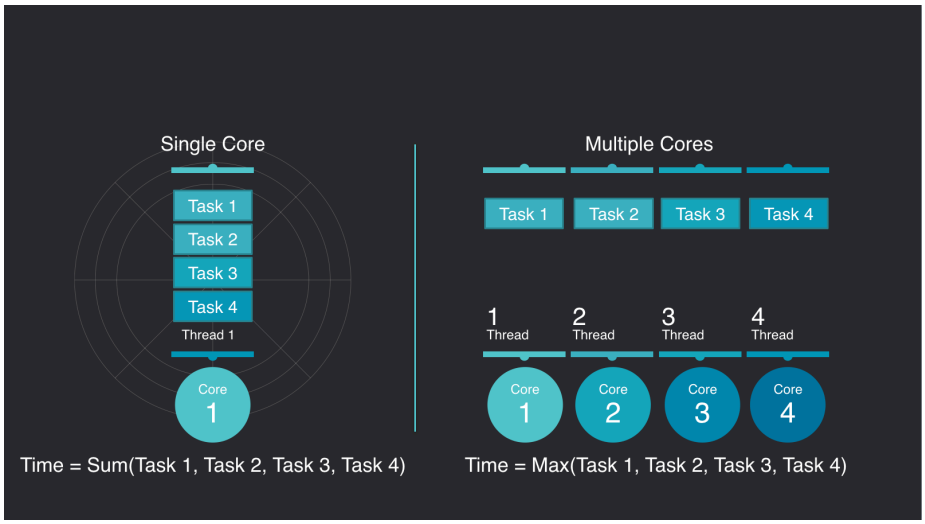

图 3:

另一个瓶颈是内存及网络操作开销的瓶颈，基本上所有的区块链项目在完成区块打包及共识的过程进行了大量的内存及网络操作，内存及网络操作比起计算操作要更加消耗资源，下图有一个简单的对比，这也极大的拖累了性能。

From Pin-down cache: a virtual memory management technique for zero-

[2]https://www.whiteblock.io/library/eos-test-report.pdf

[3]https://medium.com/coinmonks/thats-how-we-broke-jungle-testnet-twice-in-a-row-eleven-days-before-1-0-eosio-release-ef996495df

[4]https://github.com/EOSIO/eos/issues/4357

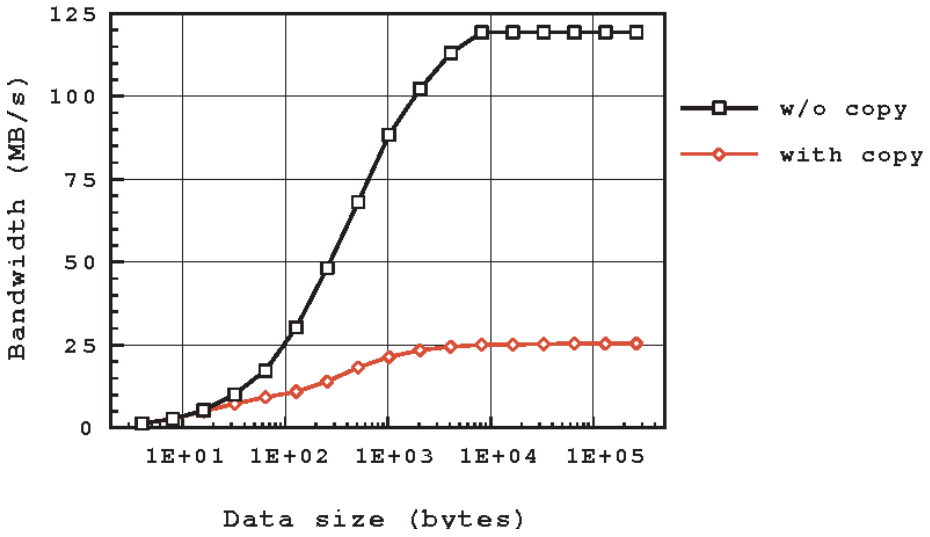

图 4:

copy communication[5]

我们深入研究并分析了这些项目的性能瓶颈，并在 POSE 中重点进行突破，以真正实现产业所需的性能。

充分发挥多核性能优势

正如前面分析所述，现代的计算体系多核已经成为主流、高性能计算已经迈入百核量级，充分发挥多核并发计算至少可以为区块链的性能带来一到两个数量级的提升。我们在分析了众多项目的技术和瓶颈后认为，以下几个主要因素是影响区块链性能的最主要原因，下面是问题的具体描述和解决方案以及 POSE 对应的实现。

首先，在多核环境下一个非常大的挑战就是如何降低引入锁和多核同步的开销。以著名的服务器软件 Apache 和 Nginx 为例，Nginx 在设计初始就极力避免多核同步的开销，因此，如下图所示，Nginx 性能上可以极大的领先于 Apache，POSE 也同样将极力避免此类开销以提供更好的性能表现。

From Apache Vs Nginx Vs Lighttpd: Comparing Performance, Resource Usage And Features[6]

其次，在现代多核 CPU 架构中有着大量不同级别的硬件级缓存，这些 CPU

[5] https://ieeexplore.ieee.org/document/669932

[6] https://iwf1.com/apache-vs-nginx-vs-lighttpd-comparing-performance-resource-usage-and-features/2

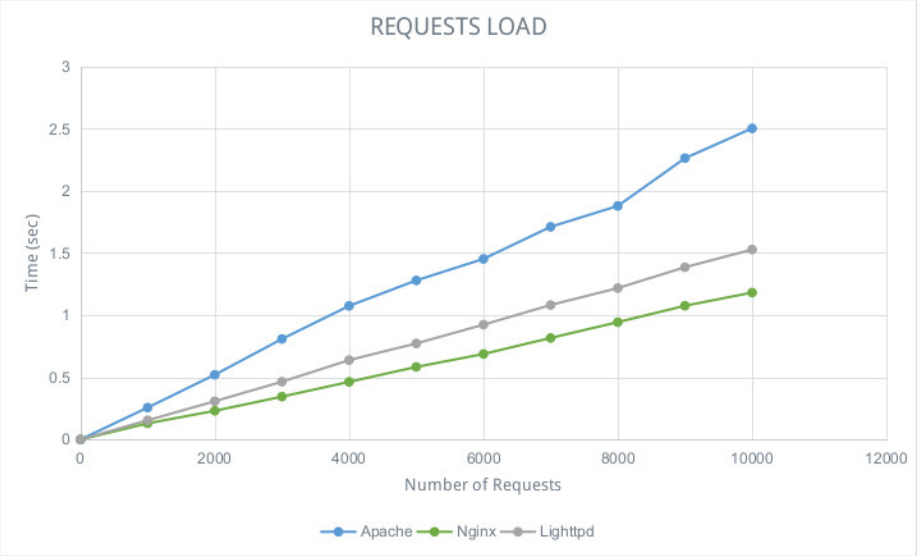

图 5:

内部硬件级别的缓存中经常被提到的是 L1 和 L2 缓存，实际上还有很多寄存器级别的缓存。在日常应用中我们并不会感知到这些缓存，但如下图所示，在对性能要求非常苛刻的情况下，这些硬件级别的缓存会严重阻碍多核性能的发挥。过去我们解决这些问题只能依靠同步机制，而现代 CPU 已经开始提供 CPU 级别的原子操作来帮助更高效的解决这个问题，POSE 将充分利用这一特性来实现更高的性能表现。

最后，需要继续发挥无锁数据结构的优势，无锁数据结构有着优秀的性能表现这个已经是业界的共识，POSE 要做的是在多核环境下仍然能够充分发挥无锁数据结构的性能优势。

From Lock-Based vs Lock-Free Concurrent Algorithms[7]

但这些并不容易。

以无锁架构为例，Bitshares 从 LMAX 中借鉴了一个高效的无锁环形数据结构[8]，原理如下图，这是一个非常好的思路，可以帮助其降低对会严重影响性能的计算机内核中的锁和信号量等同步机制的使用。但这和多核并行是一对很难处理的矛盾，这也是 EOS 摆脱单线程的很大障碍。

From The LMAX Architecture[9]

[7] https://mechanical-sympathy.blogspot.com/2013/08/lock-based-vs-lock-free-concurrent.html

[8] https://martinfowler.com/articles/lmax.html

[9] https://martinfowler.com/articles/lmax.html

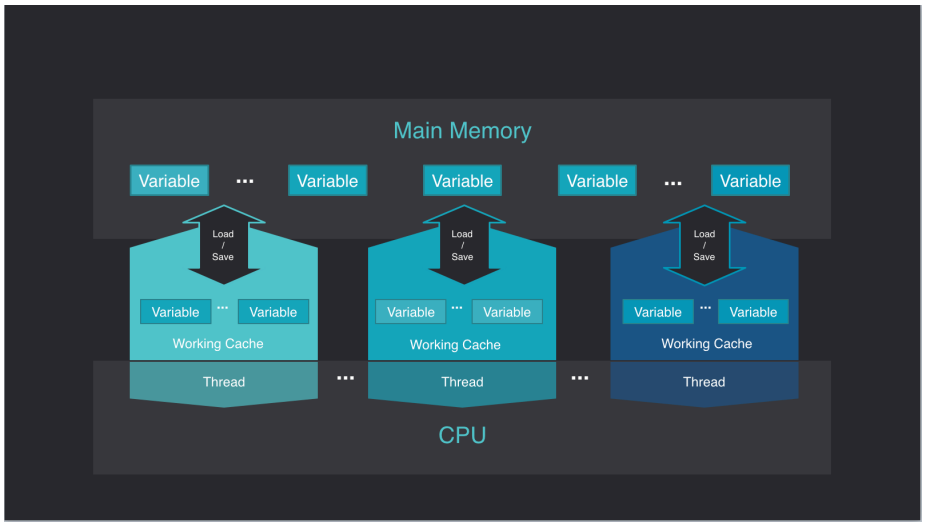
Main Memory
Variable
Variable
Variable
Variable
Variable
Load / Save
Load / Save
Load / Save
Variable
Variable
Variable
Variable
Variable
Variable
Working Cache
Working Cache
Working Cache
Thread
Thread
Thread
CPU

图 6:

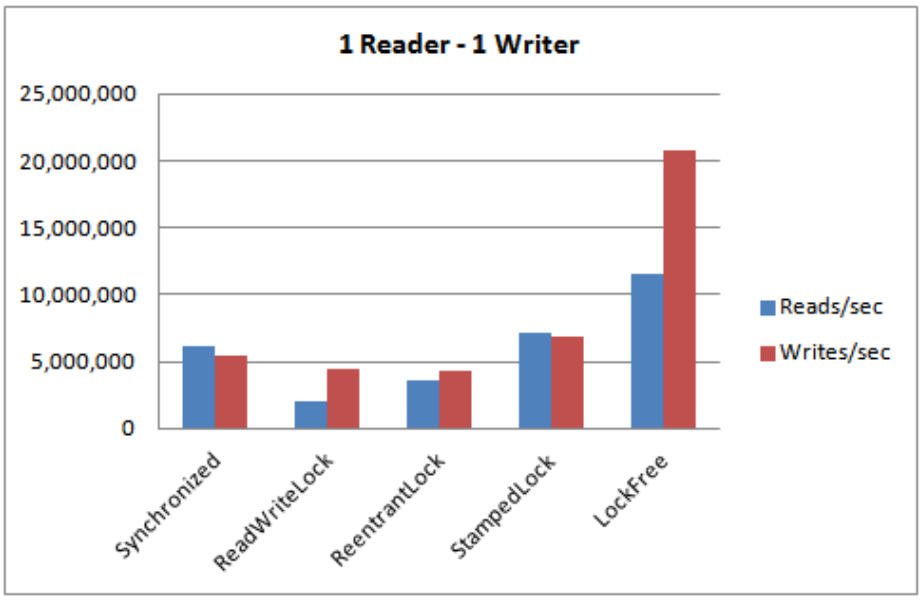
1 Reader - 1 Writer
25,000,000
20,000,000
15,000,000
10,000,000
5,000,000
0
Synchronized
ReadWriteLock
ReentrantLock
StampedLock
LockFree
Reads/sec
Writes/sec

图 7:

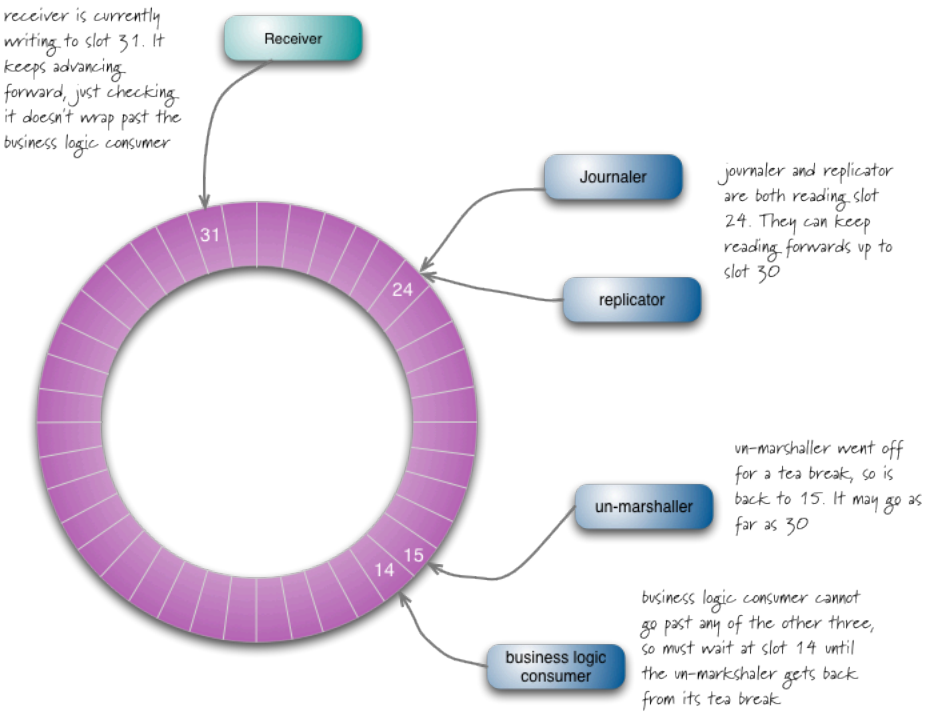

图 8:

研究并总结了这些因素后，POSE 将实现一套多核环境下的无锁智能合约执行架构以实现并发，这将有别于目前所有区块链项目的智能合约执行体系。在这个体系里，合约执行仍然分成七层[10]，依靠 Shard 实现并发，如下图所示。

```
Block
    Region
      Cycles (sequential)
        Shards (parallel)
          Transactions (sequential)
            Actions (sequential)
              Receiver and Notified Accounts (parallel)
```

在此基础上，POSE 将在代码层面实现一套全新的多核智能合约执行环境来实现无锁智能合约执行架构。这个架构主要分为两个部分，第一部分是关于数据分片，链上数据将根据 Hash 进行分片，交易根据引用的数据所属分片

[10]https://github.com/EOSIO/Documentation/blob/master/TechnicalWhitePaper.md#minimizing-communication-latency

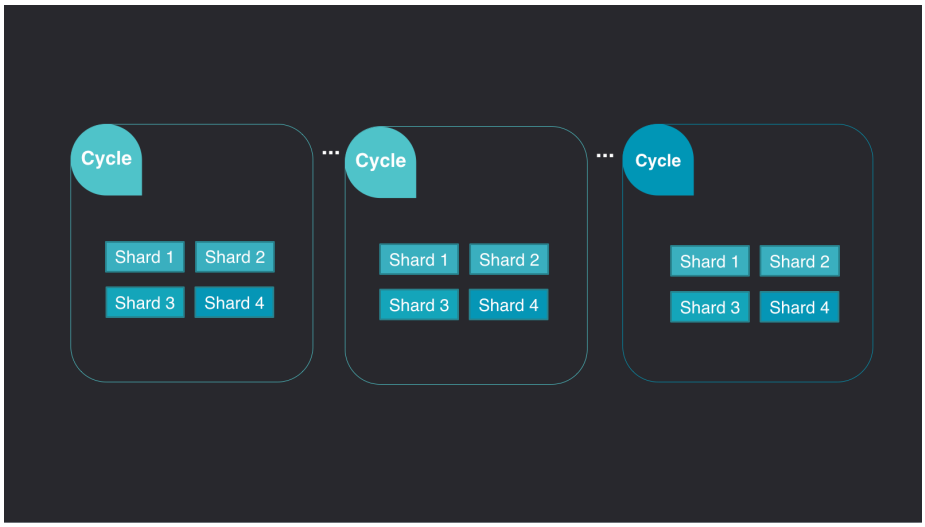

图 9：

进行分类，每个区块中的交易根据分类进行并发，最终结果在并发执行后聚合打包，分片并发的逻辑过程如下图所示。

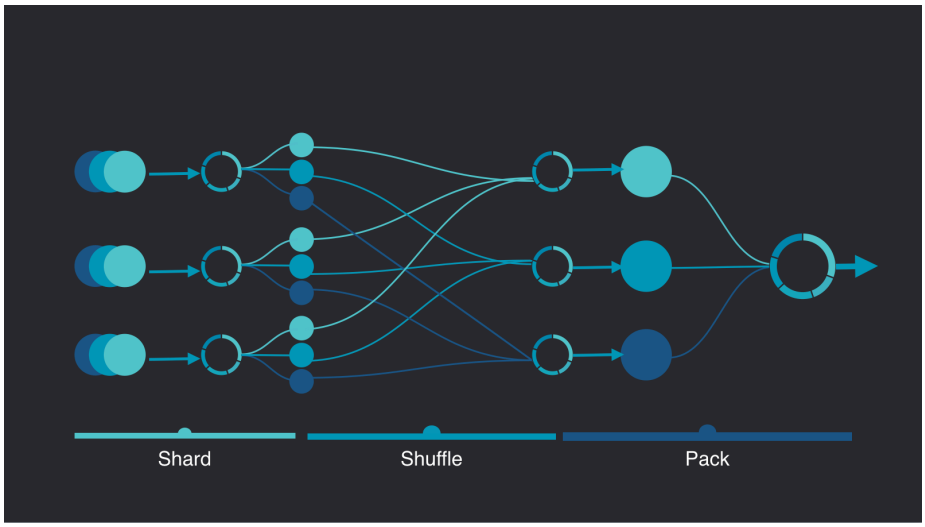

图 10：

第二部分是执行智能合约的虚拟机，需要让执行智能合约的虚拟机基于新的数据分片模型在多核上并发运行，这里需要实现一个多核间共享的无锁数据访问空间，在这个空间里完成数据分片，同时还需实现一个可以在多核上并发运行的虚拟机引擎以及配套的数据分片总线，整体结构如下图所示。

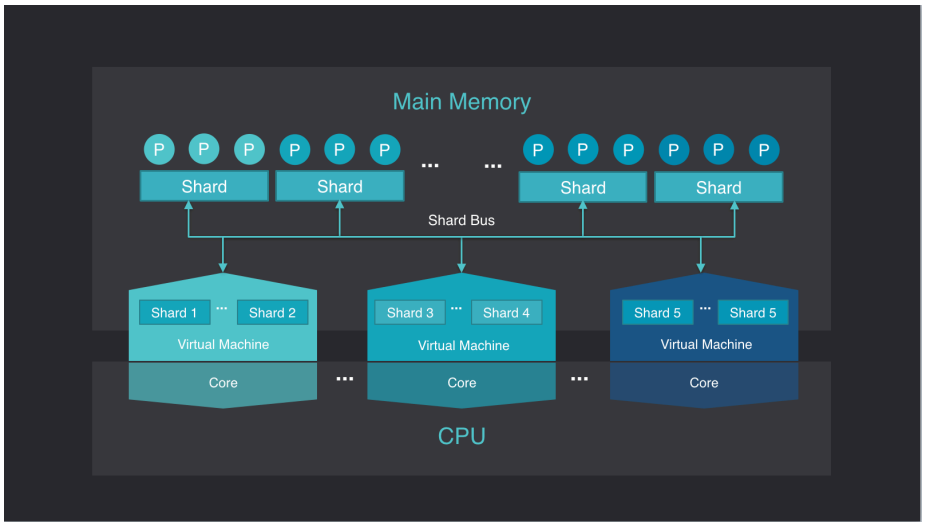

图 11:

综合这两部分，POSE 将能够将 TPS 实现数十倍到百倍左右的提升，这虽然已经足以成为下一个里程碑式的提升，但对于 POSE 来说仍然不够。这就需要引入另一个技术，即内存零拷贝技术。

引入内存零拷贝技术

在前面分析区块链的性能瓶颈时，我们提到了另一个性能瓶颈，内存和网络操作所带来的开销。下图是一个经典的程序运行过程，这里可以看到，仅从设备层到内核层到用户层，就有着多次的内存拷贝。

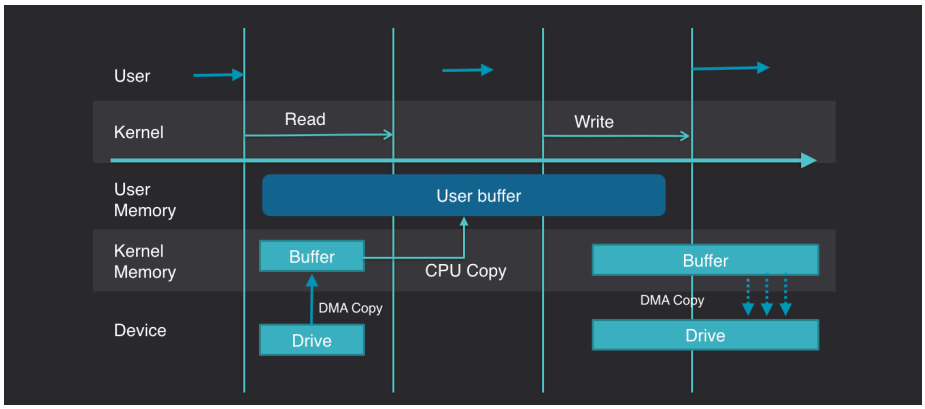

图 12:

而实际上，程序中的内存拷贝远不止这些，目前的区块链项目中普遍有着大

量的临时数据、缓存数据、持久化数据等数据空间之间的内存拷贝。这还将引发另一个问题即大量的内存碎片，会进一步加剧性能的损失。在计算机体系结构里，内存操作一向是是非常耗时的，根据研究，磁盘及内存的访问速度发展远远没有跟上计算能力的发展速度[11]。因此 POSE 中必须引入一个新的机制来解决这个问题。

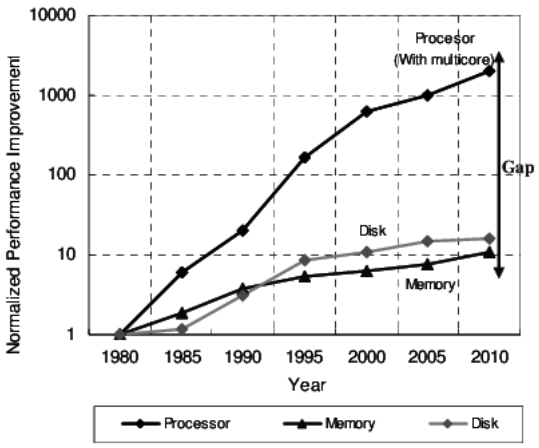

图 13:

From A layout-aware optimization strategy for collective I/O[12]

问题的解决主要从两方面实施，一是充分利用 COW（copy on write）技术，COW 在区块链的数据存储上尤其在 Merkle Tree 上应用非常广泛也非常成功，但在内存处理上却并没有充分应用。COW 的原理如下图所示，仅当数据修改时才需要进行内存拷贝，在内存中充分利用 COW 技术将极大提升性能。

POSE 将使用 COW 技术大幅改进现有的内存操作，将易变的状态和不易变的数据进行分离，减少不必要的内存拷贝，降低内存的碎片化，提高吞吐性能。

二是网络及磁盘等 IO 等操作的零拷贝，通过使用更高级的系统中断回调，在网络操作和磁盘读写时直接使用同一块内存缓冲，从而实现在多个系统态之间的内存零拷贝。

这里，POSE 还将实现一个新的内存数据库架构，让内存数据库同样减少内部内存拷贝操作和磁盘读写拷贝。依托于这些实现，POSE 的性能将进一步获得指数级的提升，从而真正意义上达到工业级的性能。

[11] https://www.researchgate.net/figure/Performance-improvement-trend-of-processor-memory-and-disk_fig1_220717222

[12] https://dl.acm.org/citation.cfm?doid=1851476.1851530

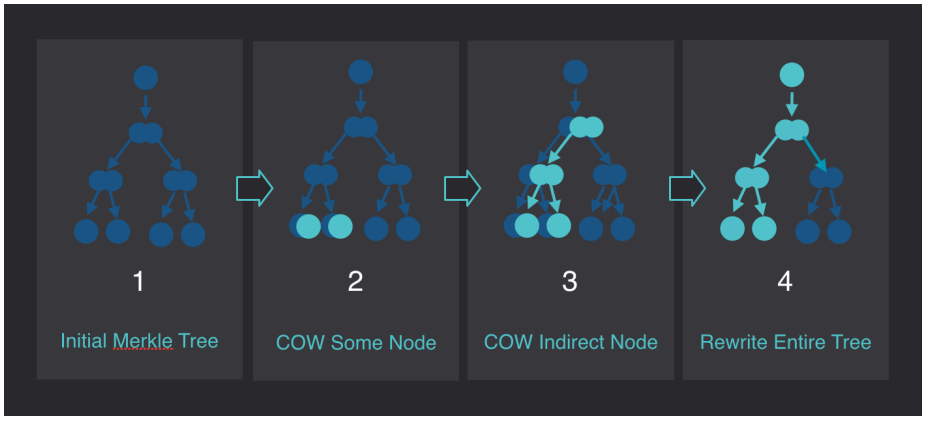
1
Initial Merkle Tree
2
COW Some Node
3
COW Indirect Node
4
Rewrite Entire Tree

图 14:

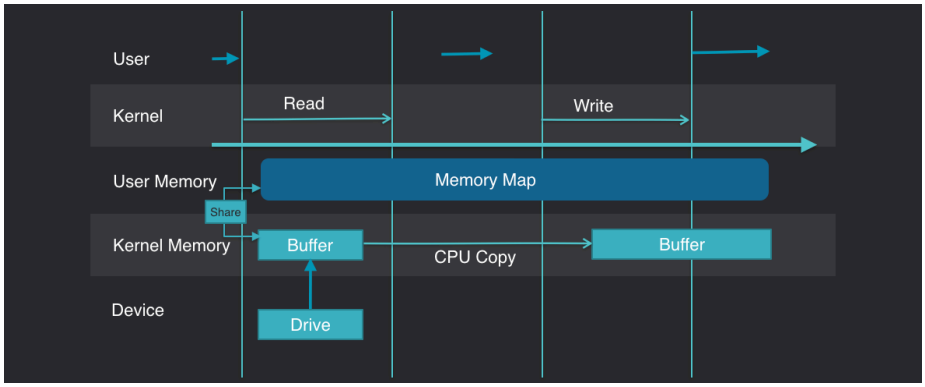
User
Kernel
Read
Write
User Memory
Memory Map
Share
Kernel Memory
Buffer
CPU Copy
Buffer
Device
Drive

图 15:

确定性，降低产业链上运行成本

产业的发展有着和摩尔定律非常类似的变化，即单位成本呈指数下降，但对数据量和处理能力的需求呈指数形式上升。前面我们讲了如何支撑产业级的数据量和处理能力，为了满足指数级降低成本的需求，POSE 将从确定性角度出发，通过以下几方面让广泛的产业在链上真正低成本的运行。

高效治理机制

POSE 将从技术上实现一个高效的投票治理机制，包括投票选举和投票决议两部分。

首先，POSE 将实现一个去中心化的投票功能，所有参与 POSE 的用户都拥有投票权力。用户通过一个系统内置的智能合约进行投票，选举和决议则根据限定时间段内的投票情况自动完成。

节点选举

参与 POSE 的用户可以随时投票，投票结果实时统计，定时根据投票结果选举出主节点和备用节点。主节点负责保障整个 POSE 网络的稳定运行，共同负责交易的验证、合约的执行以及区块的生成。备用节点负责在主节点出现故障时代替出现故障的主节点，继续保证网络的稳定。

决议

POSE 会从技术上实现一个议题提议和决议的功能，当一个议题被提出时，首先将在基金会成员中投票，当出现分歧时会扩大到所有主节点投票，然后会扩大到所有备用节点投票，最后会扩大到整个社群投票直至分歧解决。

快速确认

得益于主节点的数量有限，每个节点的连续出块数量大大减少，交易可以快速确认。这将极大的降低企业链上运行的等待成本。

利用侧链技术，企业可以搭建与 POSE 底层结构一致的侧链，直接与主链进行交互。根据不同需求，在不同侧链之间合理分配工作量，从而实现 TPS 整体提升，进一步降低营运成本，大幅提高灵活性。

杜绝分叉

得益于主节点和备用节点的设计，POSE 从概率上可以得到一个极低概率的分叉可能性，故而从很大程度上可认为基本足以杜绝分叉。

通过计算节点之间的延迟，连接延迟较低的节点之间可以相邻出块。这种算法设计，对比以前节点随机排序，大幅提高系统出块效率，并且有效降低软分叉的风险。这将有效的为企业节约大量处理分叉容错的成本。

动态费率

为了维护网络的稳定 POSE 将设计一个动态费率机制，避免网络出现极度繁忙的情况，帮助企业实现链上运行成本的可控性。

中本聪曾经说过，“矿工应该发现遵守规则会比破坏系统和自己财富的有效性更有利可图”。适量手续费的存在，可以更好的激励生产节点打包区块，促进整个 POSE 生态网络良性运转，互惠互利平稳运行。

费率的动态调整设计，也使得整个服务更加有弹性，企业可以根据自己的服务合理量化自己的经营成本。

成本隔离技术

让所有主节点分担企业的运行成本是不合理的，因此 POSE 将实现一套成本隔离技术，让不同企业间的计算成本实现隔离，让企业能够确定性的承担自己业务的成本。

结合动态费率机制，链上运行的企业将不再会受到其他企业的干扰。有效降低企业因为价格波动等因素带来的对于链上运行成本不可控的担忧。同时也将一定程度降低企里链上运行的成本，帮助企业更好的发展。

便捷丰富的开发扩展

产业级智能合约引擎

满足产业复杂的多样的需求，我们需要一个强大的智能合约引擎。这里，Wasm(WebAssembly) 因为诸多优点已经广泛被认为是 EVM 的理想替代品。POSE 除了继续沿用 Wasm 作为智能合约引擎外，还将引入一个 QuickJS 作为下一代智能合约引擎的一个强大补充。

QuickJS 作为最小的可嵌入的 Javascript 引擎，完整支持 ES2019 规范，被认为是目前最优秀的 Javascript 引擎。它的加载速度非常快，根据评测可以到百微秒级别，远低于 Wasm 的 30 甚至数百毫秒的量级。而 Javascript 是目前应用最广泛的脚本语言，有着非常广泛的应用基础。采用 QuickJS 作为下一代智能合约引擎的强大补充必将大大增强 POSE 的开发友好度。

而 Wasm 是由 W3C 工作组开发的 Web 浏览器的标准，包括 Google，Apple，Microsoft，Facebook，Mozilla 等众多厂商在为其做贡献，它设计的初衷便是使代码可以在任何浏览器中部署，并且实际上也达到了这个效果。另一方面，Wasm 性能也非常高，Wasm 的编译目标是尽可能接近原生机器代码，同时仍然与平台无关。这些都有助于将小型二进制文件通过互联网分发到不同设备。此外，Wasm 已经有很多年的历史，包括编译器和标准化团队。这些都将让 POSE 的智能合约引擎更加强大。

POSE 的这个智能合约引擎还需要实现以下几点：

- 内存安全，智能合约沙箱以及平台无关。
- 支持 64 位和 32 位整数运算，与 CPU 指令一一映射。去除浮点运算，实现确定性。
- 基于编译器优化，将智能合约编译成二进制文件进行分发和执行。
- 智能合约分片和高效的虚拟机加载。

主流开发语言的支持

基于这样一个智能合约引擎，POSE 将有着最广泛的开发语言支持，包括 C、C++、Go、Java、Javascript、Python、Rust，甚至 Kotlin 等非常新的语言，这意味着开发者可以使用熟悉的任何语言编写智能合约。

POSE 还将提供一系列标准的合约模板和帮助，甚至不同语言间的智能合约转换工具，进一步帮助不同语言的开发者在 POSE 上实现开发。

结束语

跟随路线图，POSE 将为人类进入新一代区块链乃至新一代互联网提供一个更真实的可能。我们将会在未来一到两年实现这一愿景，在下一个时代路口，POSE 将携手所有致力于区块链领域应用的企业和人，一起迈进新时代。我们足够幸运，生活在一个充满难以想象的先进技术和新创意的创新时代，使我们得以站在巨人的肩膀之上不断努力。