The Future of Fintech

By Andrew Bakst

March 3, 2020

The Future of Fintech

Consumer and enterprise financial applications will become the first mainstream products built on public blockchain infrastructure. Current financial infrastructure is decades old but has been immune to disruption due to expansive regulatory and network moats. These moats have afforded finance firms some of the largest market capitalizations in the world: Wells Fargo at $200bn (regulatory moat), Visa at $500bn (network moat), the list goes on.

There are four layers to finance's current stack:

- the vaults (commercial banks, such as Wells Fargo),
- the rails that connect the vaults (back-end payment networks, such as VisaNet),
- the APIs that provide access to either the rails (payment processors, such as Stripe and Square) or the vaults' data (data services, such as Plaid),
- and the digital wallets that build on top of these third-party APIs (such as Venmo).

Certain firms hold positions in multiple layers of the stack. Visa owns VisaNet (rails) and Plaid (data APIs). Amex is a card issuer (vault) and owns a proprietary payment network (rails). Square builds APIs that give access to the ra-ils *and* forces users to use their application (the Cash app) to gain the benefits of those APIs. The entire stack will be disrupted by the triple accounting technology known as public blockchain infrastructure.

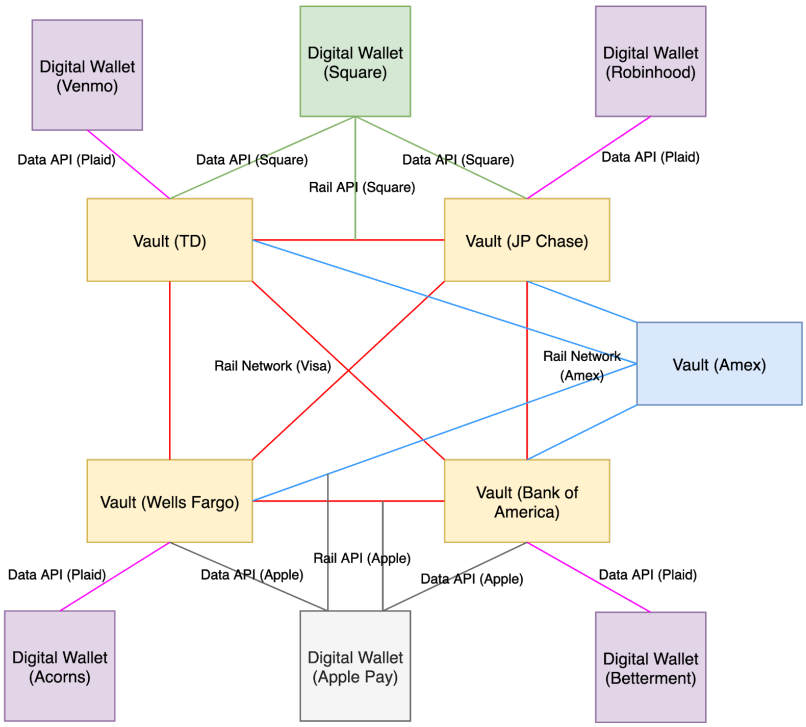


Image 1: Current fintech infrastructure

Public blockchains will disrupt each layer of the stack to varying degrees:

- Public blockchains allow for the vault and the rails to exist at the same layer, the blockchain itself, which allows for full-scale disruption of the rail layer.

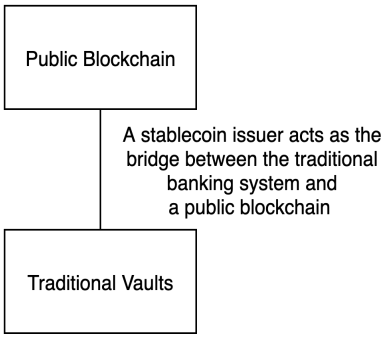
- Public blockchains open-source the code necessary to transact on the blockchain and return access control of user data to the users themselves, the combination of which allows for full-scale disruption of the API layer.

Thus, the rail and API layers will be disrupted fully, so long as the vault layer achieves the minimum viable disruption of moving fiat currency onto a blockchain (Mainstream users will still wish to primarily transact with and lend stable fiat currencies, not volatile cryptocurrencies such as Bitcoin or Ether.). The process of moving fiat currency onto a blockchain is already underway due to the increases in efficiency that the technology provides over today's current banking infrastructure. However, the methods of moving fiat currency onto a blockchain vary significantly, with each method posing a different degree of disruption at the vault layer.

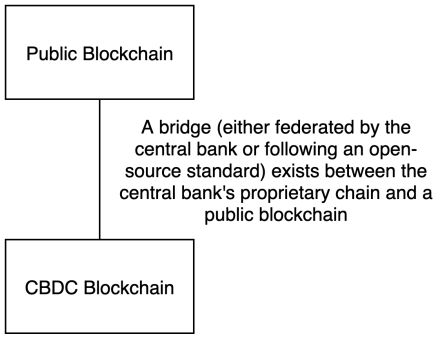
The Future of Vaults

The vault layer's disruption has three possible degrees, dependent on the method of 'fiat transitioning' used (Once fiat currency is moved onto a blockchain, it is termed a 'stablecoin.'):

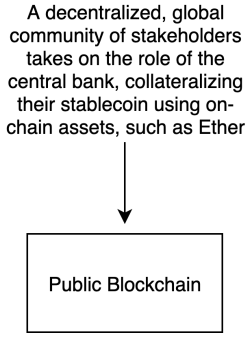
Minimal viable disruption occurs when a company issues fiat-denominated stablecoins on a public blockchain, with each stablecoin backed 1:1 by that issuer's deposits at a commercial bank. These stablecoins are known as **dollar backed stablecoins (DBSs).**



Medium scale disruption occurs when central banks issue their currencies directly on a blockchain (either on their own proprietary blockchain or as an asset on a public blockchain--the latter being unlikely, at least initially). These stablecoins are known as **central bank issued digital currencies (CBDCs).**



Maximum viable disruption occurs when fiat-denominated stablecoins become backed by on-chain assets at collateralization ratios deemed to be sufficient by a decentralized community of shareholders (a decentralized, global central bank). These stablecoins are known as **fully collateralized on-chain stablecoins (FCOSs).**



Trust Dependencies

These three stablecoin solutions differ significantly in their trust dependencies, dependencies which directly impact the competition generated within the banking system by the solution (as shown in the following **Competition → Efficiency** subsection). For reference, current holders of digital fiat currency need to trust both that currency's central bank (namely that the central bank will not inflate away the value of the currency) and that central bank's corresponding commercial banking system (such that their deposit credits at commercial banks are always fully redeemable).

DBSs require users to trust both the traditional banking system and the issuer of the stablecoin. While the issuer of a stablecoin could be the same entity as the vault where the issuer's deposits are held (which would equate the trust dependencies of DBSs to the those of the traditional banking system), commercial banks have veered away from this strategy, instead focusing on building private blockchains for intra and inter-bank settlements. Commercial banks have likely shied away from launching dollar-backed stablecoins due to the incumbent's dilemma: the system works well for them as is; switching to a blockchain-based currency would eliminate their monopoly on inter-bank transfer fees, which they capture in every payment.

CBDCs reduce (with the potential to eliminate) trust in modern commercial banks but maintain trust in central banks (The term 'modern commercial banks' is used to to denote federally-insured commercial banks, the global standard.). CBDCs allow users to custody their own dollars electronically: whereas the modern banking entails digital credit, CBDCs entail digital cash. Thus, users have the option of removing themselves from the modern commercial banking system at scale, through either engaging in lending directly on public blockchains (becoming their own bank) or lending their funds only to commercial banks that use free-market insurance solutions (purposefully ignoring federally-insured commercial banks). For trust dependencies on modern commercial banks to be fully eliminated, all holders of a CBDC would need to exit via one of the two methods described.[1]

FCOSs eliminate trust dependencies in both modern commercial and central banks, shifting trust toward a decentralized network of decision makers. Anyone can join this network of decision makers by owning the FCOS's governance token.[2] Holders of the FCOS's governance token make the same decisions that central banks make with regards to collateral requirements (what assets can be used as collateral, what is an appropriate collateralization rate for these assets) and interest rates (what is the cost of taking a loan from us, how much do interest do we pay on excess reserves), as well an additional decision that central banks do not make: how to establish a decentralized, resilient oracle network.[3]

---

[1] Even if just one federally-insured commercial bank fails and the central bank's insurance solution cannot fully cover the difference, net losses would likely be amortized across all CBDC holders (as they have been in the past).

[2] Each FCOS has a governance token. This is known as a duel token system: for example, MakerDAO calls its governance token 'Maker' and its FCOS 'Dai'.

[3] Oracles are computers/people that input data into blockchains. FCOSs require them to maintain accurate fiat-denominated prices of their collateralized assets, such that they can liquidate any position whose collateralization falls under the encoded collateralization ratio.

Competition → Efficiency

Of note, all of the solutions increase competition between central banks that issue fiat currencies: the open nature of blockchain-based money renders foreign exchange significantly easier to conduct,[4] while the bearer nature of blockchain-based money highly decreases the risk of confiscation of 'foreign fiat' by a local's native government (which has happened frequently in the past to citizens of the developing world). However, the set of trust dependencies described above portend whether the solution generates additional newfound efficiencies within the banking system, under the assumption that competition leads to efficiency

DBSs, due to their increase in trust dependencies, do not create any additional efficiencies in the banking system. There may be marginal efficiencies created by shifting power in commercial banking to the vaults used by the most successful DBSs, but these marginal efficiencies would seldom be felt by users.

Unlike DBSs, **CBDCs** have the potential to further competition in commercial banking by allowing individuals to either be their own bank or use neo-commercial banks insured by free markets. Whether competition ultimately exists between commercial banking solutions (federally-insured vs. self-insured vs. free-market-insured) will be determined by consumer preferences and central banks' wherewithal to remove their monopoly on banking insurance solutions.[5]

FCOSs give users a viable exit from modern commercial and central banking.[6] They are the first stable currency to do so; other exit options, such as Bitcoin or Ether, are not stable. Due to their existence outside of the traditional banking system, FCOSs can only be lent with the advent of free market insurance providers.[7] Consequently, FCOSs have the potential to pressure central banks into enforcing that commercial banks adopt free market insurance solutions, should consumers signal their desire to hold a currency unable to be devalued through commercial bank malpractice (Commercial bank malpractice is often amortized across all corresponding fiat holders, as governments historically bail out their failed banks.). FCOSs also have the potential to pressure central banks to discontinue the devaluation of their currencies (measured through the decrease of treasury bond yields), as consumer interest in FCOSs could also signal

---

[4] Foreign exchange is currently dominated primarily by dark pool exchanges, which, due to their opacity, render foreign exchange extremely difficult for mainstream users to conduct. Users often pay exorbitant fees while needing to be physically present to conduct the transaction. Crypto exchanges are gradually becoming foreign exchange markets, with users obtaining fair prices from their computers.

[5] Consumers should be excited by the concept of free-market insured banks, as more competitive banking would entail better interest rates on deposits along with significantly less risk of a widespread banking failure. Central banks should be excited by this as well, as it reduces the burden on them to both bailout banks and define appropriate bank reserve requirements. Instead, insurance providers could analyze a bank's outstanding loans and offer competitive revenue-share based premia. Their bids would also likely include their own reserve requirements, also calculated via in-house analysis. The insurance provider could choose to not insure (or veto, depending on the agreement between the bank and the insurer) loans that project negative expected value. Free market economies will likely be more willing to allow for this novel form of commercial banking. However, those that think they live in a free market economy may not actually live in a free market economy.

[6] The decentralized banks behind FCOSs use over-collateralized loans to ensure stability of the system. Should the value of these over-collateralized loans crash and an FCOS become insolvent, the FCOS's governance token holders are punished (as the FCOS's smart contract prints and sells a potentially infinite amount of their governance token, in order to obtain the funds to become fully solvent). Government-backed central banks have no self punishment mechanism for the insolvency of the systems they perpetuate.

[7] Which do not currently exist, and thus FCOSs cannot currently be lent fractionally.

mainstream desire to hold a currency unable to be devalued by central banking policies.[8] Lastly, FCOSs are inherently self-competitive: their stakeholders can always exit via hard fork, creating a new decentralized central bank with a new governance token (and a new FCOS) that follows different risk parameters from the original. A hard fork can be executed with little technical overhead: the new FCOS could copy the old FCOS's code nearly verbatim.

Theory into Reality

The degree to which each solution is adopted is not only a byproduct of trust dependencies and these dependencies' corresponding efficiency improvements, but also a factor of implementation difficulty. Technically, each solution suffers little implementation overhead, as public blockchains and smart contracts have been under development for approximately ten and five years, respectively. The largest implementation difficulties entail those of distribution (DBSs), reputation risk (CBDCs), and the need for new entrants (FCOSs):

- Stablecoin issuers must find and convert users for whom holding a **DBS** is a magnitude order of improvement over their current banking options.
- Central bank employees must convince themselves that creating a **CBDC** is worth the career risk of failure.
- Insurance providers must be convinced to expand their offerings, such that **FCOSs** can be lent fractionally by neo-commercial banks.[9]

	Dollar Backed Stablecoins	Central Bank Issued Digital Currencies	Fully Collateralized On-Chain Stablecoins
Trust Dependencies	Central banks and their corresponding commercial banking system; stablecoin issuers	Central banks and (potentially) their corresponding commercial banking system	Decentralized network of stakeholders
Efficiency Improvements	Competition between central banks increases; little to no change at the commercial banking level	Competition between central banks increases; competition between commercial banks has the potential to increase	Competition between central banks increases significantly; competition between commercial banks increases significantly
Largest Implementation Challenge	Distribution: how do we get our DBS into the hands of those for whom a DBS is a significant improvement over today's options?	Reputational risk: the current system works fine; what if we implement this and it's a disaster?	New entrants: who will offer insurance for deposits of a non-government backed banking system?

It will take a significant amount of time for one of these three solutions to emerge as the ultimate winner in the pareto distribution of stablecoin solutions. It is likely that all three will thrive in

[8] Even though FCOSs are currently pegged to central bank denominated currencies, an FCOS can always break its peg should traditional central bank currencies experience inflation that significantly devalues them, which could cause a large entrance into FCOSs.

[9] Insurance providers would likely only enter the space if an FCOS achieved sufficient distribution. Because FCOS's value proposition is so unique, their distribution likely cannot be achieved through traditional advertising channels, but rather through broad consumer distrust of commercial and central banking, which lies largely outside of an FCOS's control. FCOS communities are best positioned to succeed on distribution if they educate others on the pernicious features of modern commercial and central banking, such as the Cantillon Effect.

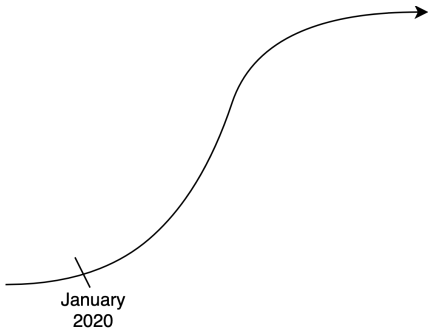
tandem for the foreseeable future, as all three provide significant advantages over the current market leader, the incumbent banking system. DBSs and FCOSs exist today (with more than $5bn and $100mn stored in each, respectively). CBCDs will be widespread within two years (possibly sooner, as China is prepared to launch their solution this year). All three solutions have the potential to benefit from exponential growth due to an awaiting flywheel effect triggered across relevant parties:

- If DBSs become adopted, large commercial banks will enter the market, which will further adoption, which will further the entrance of more large banks, etc.
- If CBDCs become adopted, prominent central banks will enter the market, which will further adoption, which will further the entrance of more central banks, etc.
- If FCOSs become adopted, an insurance provider/alternative-banking firm will enter the market, which will further adoption, which will further the entrance of more insurance providers, etc.

All three solutions also benefit from the network effects of liquidity:

- As more consumers use any of the three solutions, more payment processors will offer that solution, which will increase consumer adoption, etc.[10]

These two exponential growth impeti have yet to be triggered, leading to the assumption that, in terms of adoption, we are likely at this point of the S-curve:



Triggering the S-Curve

Currencies serve three primary purposes: to save (store of value), to pay (medium of exchange), and to account (unit of account). There are three types of users who will be most likely to switch to one of the aforementioned currency solutions, one because of blockchain-based currencies' ability to enhance their ability to save, and the others because of blockchain-based currencies' ability to improve the efficiency of their payments:

- The savers: Citizens of developing nations who do not trust their central banks due to past or current abuse. There are numerous nations with currencies suffering from annual inflation of over 10%.
- The payers: Businesses that conduct a significant amount of international transactions, as illustrated in more detail in the following **Future of Payments** section.

[10] However, liquidity is not a defensible moat for one form of stablecoin against its competitors, as each stablecoin can be exchanged frictionlessly using chain-native automated market makers, such as Uniswap.

- The payers: Individuals (specifically migrant workers from developing nations) that conduct a significant amount of international transactions, as illustrated in more detail in the following **Future of Payments** section.

Certain impeti could also trigger adoption sooner than expected:

- Governments (such as China's CPC) could enforce the use of CBDCs by Chinese businesses;
- Broad central banking failure could significantly bolster the value proposition of FCOSs;
- Negative interest rates could drive citizens of developed nations to adopt DBSs or FCOSs, due to their ability to earn positive interest rates with these forms of currency.[11]

While DBSs currently account for 90% of blockchain-based currency market share, the adoption cycle is still nascent, as $5bn is infinitesimal compared to future market size. Due to governments' ability to enforce CBDCs by law and the unique value proposition of FCOSs, both of these solutions have the ability to achieve viral adoption at almost any point in time.

No matter which blockchain-based solutions succeed at the banking layer, the rail layer will be fully disrupted.

The Future of Payments

Public blockchains merge the rail layer into the vault layer, significantly simplifying the current complexity of payment networks. Current debit/credit transactions typically require four middlemen: the payment processor, the issuing bank, the payment network, and acquiring the bank (along with nine steps of interactions between these parties). International bank wires typically require four as well: the sender's bank, the receiver's bank, and two correspondent banks (along with four steps of interactions between these parties). Blockchains reduce the intermediaries in modern payment rails from approximately four to one, the chain itself. In turn, this reduces payment processes that previously took four to nine steps down to two: the payer signs the transaction and the blockchain executes the transaction.[12]

However, public blockchains are not currently adequate for all types of payments. There are three properties of a sufficient payment scheme: **scalability**, **privacy**, and **security**. Blockchain-based payments are currently unable to fully optimize these three variables simultaneously, although they will be able to in the near future, using a combination of layer-one and layer-two solutions.[13]

[11] Thus far, developed nations' consumers have signalled that earning 4-7% additional interest on savings is not worth leaving the modern banking system. However, due to the pain of losing money, this may change should bank deposits begin implementing negative interest rates.

[12] The receiver's wallet is notified via open APIs (either through light clients or a query protocol) that the payment has arrived. This was not counted as a step, as the process of web front-ends using a bank's internal API to update an account's status was also not counted.

[13] Layer-one refers to a public blockchain network itself, such as Ethereum, whereas layer-two refers to a network built on top of a layer-one chain, such as a rollup.

Scalability

Scalability can be measured via two components: clearing/settlement speed and transaction throughput. Today, blockchains can clear and settle payments significantly faster than legacy rails.[14] However, blockchains' transaction throughput is currently limited by two parameters that blockchains optimize for security rather than scalability: block space and block time.

- As block space decreases, blockchains become more decentralized (more nodes can validate the chain), which increases a blockchains' security. However, smaller block spaces limit how many transactions can be processed in one block.[15]
- As block time increases, validators have more time to verify the correct state of the chain (more time allows for more communication among nodes), which increases a blockchains' security. However, extended block times limit how frequently blocks (and therefore transactions) can be processed.

Consequently, blockchains' throughput is currently inadequate for most payment use cases: blockchains can process ~15 tx/sec at a fee of ~$0.11 per transaction[16]. At times of high demand to transact, fees can increase by an order of magnitude (up to $5.00), while tx/sec remains constant.[17] These transactions are not even private: they publicize pseudo-anonymous information, such as users' public key addresses, as well as how much money was sent. Pseudo-anonymous checking accounts are unacceptable for most corporations and individuals, as they would enable competitors/peers to monitor essential operations of their businesses/lives.

Private transactions are possible, but require more complexity and thus consume a larger portion of block space than regular transactions: the result is that blockchains can only process approximately two private tx/sec currently, which means that blockchains could not even handle the full demand of American bi-directional international wires today[18] (Although, even at this speed, blockchains still present an alternative solution to a market previously monopolized by banks.).

In order to make blockchains more scalable for both inter-bank wires and everyday transactions (a use case that blockchains will eventually serve), developers are in the process of deploying competing horizontal scaling[19] schemes: both layer-two solutions and sharding. Both layer-two solutions and sharding maintain similar block times and block sizes as today's standards, but

[14] Ethereum transactions clear in 13 seconds, meaning that the funds are already in the receiver's account, and settle in two minutes, meaning that the blockchain now recognized the payment as finalized. Depending on one's definition of 'clearing', one could claim blockchain-based transactions clear in under a second, as wallets can confirm that the transaction is in the mempool instantaneously (Coinbase provides this service already when depositing funds into their wallet.).

[15] This is a feature, not a bug. Public blockchains attempt to ensure that the requirements to run a node are as low as possible, furthering decentralization, which furthers security.

[16] These reference numbers relate to Ethereum. Other public chains may differ slightly.

[17] Users bid up fees to have their transactions included in the next block, but the number of transactions in the next block is unchanged.

[18] We can safely assume there are less than five international tx/sec conducted by United States banking members.

[19] Horizontal scaling can be summarized as 'adding more nodes,' whereas vertical scaling can be summarized as 'increasing block size.'

still increase throughput drastically by adding more nodes to the network and dividing the tasks of processing transactions between distinct subsets of nodes.

Layer-Two: A Brief Overview

The horizontal scaling provided by layer-two's will enable public blockchains to process pseudo-anonymous transactions on par with current rails' capabilities and private transactions more efficient than current rails' international capabilities, while maintaining fees that are still orders of magnitude lower.[20] Within two years, layer-two solutions will be able to process fully private transactions within the same realm as current rails' capabilities, so long as innovation in the zero-knowledge proof space continues to improve at rate of at least 2x (in terms of both generation time and compactness of the proof).[21] Additionally, if layer-one solutions successfully implement sharding, the combination of layer-twos on top of sharded layer-ones will allow for an order of magnitude improvement (or two) over today's current system in both speed, fees, and privacy (within two years).

There are two leading layer-two solutions: zk rollups and optimistic rollups. These two solutions share the same leading design principle: a decentralized community of nodes executes transactions outside of the layer-one blockchain, expanding block space 'out-of-protocol', but posts enough data to the layer-one chain such that layer-one can verify the correctness of the transactions (either through fraud proofs or validity proofs). This approach leverages the best feature of layer-one chains, security (which is often measured in terms of decentralization, and layer-one chains are inherently more decentralized than layer-twos), while significantly enhancing scalability.[22]

[20] Layer-two solutions move the execution of the transaction off-chain. These transactions still leverage the layer-one chain's security, as these layer-two solutions require the posting of all transaction data to the layer-one chain. Blockchains, in this case, are simply used to store data and verify computation (via validity or fraud proofs), which requires significantly less block space and computation than if transactions were to be executed on the layer-one chain.

[21] Zero-knowledge proofs have been improving at a rate faster than this, due to the immense amount of capital deployed into the hands of cryptographers working on them. The 'software-only and 'open-source' nature of zero-knowledge proof innovations has allowed innovation to compound at an astonishing rate.

[22] Of note, sharding will leverage almost an identical similar design principle, with shards executing transactions outside of the relay chain but posting enough data to the relay chain such that the relay chain can verify its correctness.

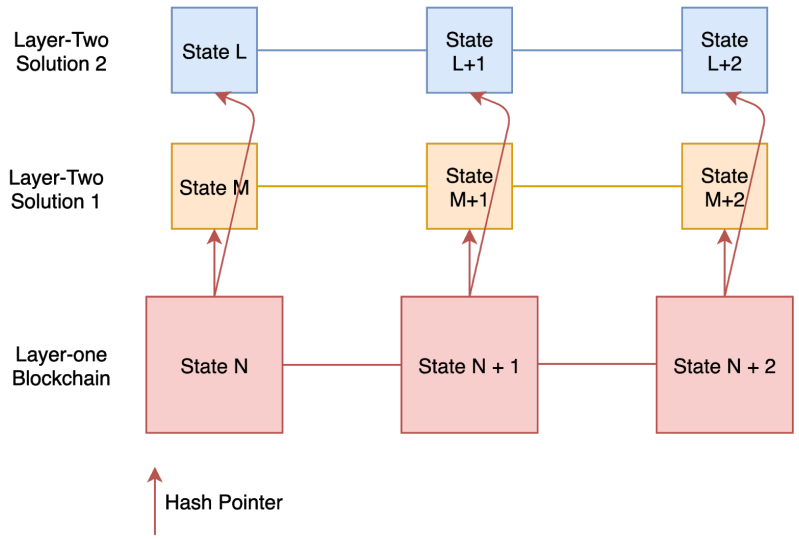


Image 2: A modern blockchain scaled through competing layer-two solutions

It requires an additional step to move funds between layer-one and layer-two chains, similar to how funds are moved between checkings and savings accounts in traditional banking. However, once funds reside in a layer-two solution, it still only requires one middleman (now the layer-two chain) and two steps to execute a transaction.[23]

Sharding: A Brief Overview

Sharding is very similar in design to layer-twos: both horizontally scale blockchains, relying on an increase of validators rather than an increase in the size required to run validators; both are a set of individual blockchains that use a base chain to shuffle their validator sets and verify the correctness of their state. Layer-two solutions can be termed as 'free market' versions of sharding due to their implementation as smart contracts on their corresponding layer-one chains: no developer or validator is forced to use or validate a layer-two solution; they must explicitly choose to, whereas developers and validators will be forced to adopt to a sharding environment should one be implemented.[24]

Sharding expands the block space of a blockchain 'in-protocol' by transitioning a blockchain of block size X into many 'sub-blockchains' of block size X (In the case of Ethereum, 64 blockchains of block size 128kB), with all of these blockchains secured by one larger, shared blockchain. Each of these 'sub-blockchains' are called shards, while the shared blockchain that secures every shard is called a 'relay chain' (also referred to as a 'beacon chain.').

[23] To move funds between layer-two solutions, two steps are required (exiting one layer-two and entering another), while sharding only requires one step to transfer funds between shards. This discrepancy in efficiency emerges because each shard monitors the entire state of the system through its hash pointers to the relay chain, whereas a layer-two solution is only concerned with its smart contract on the layer-one system (not the entire layer-one system). However, this design difference is also why sharding is taking significantly longer to develop than layer-two solutions.

[24] Developers have limited ability to exit via hard fork if they do not wish to migrate to a sharded environment, as they would likely lose interoperability between other smart contracts that make their smart contracts more powerful (The power of trustless interoperability is why blockchains have been termed 'world computers.').

Relay chains use crosslinks to confirm shard blocks, so that specified shard blocks can become finalized in the relay chain. Shard chains use hash pointers to reference relay blocks, so that shard chains can securely execute cross-shard transactions. The innovations required to make sharding work are centered around stateless clients, erasure coding, and secure networking.[25]

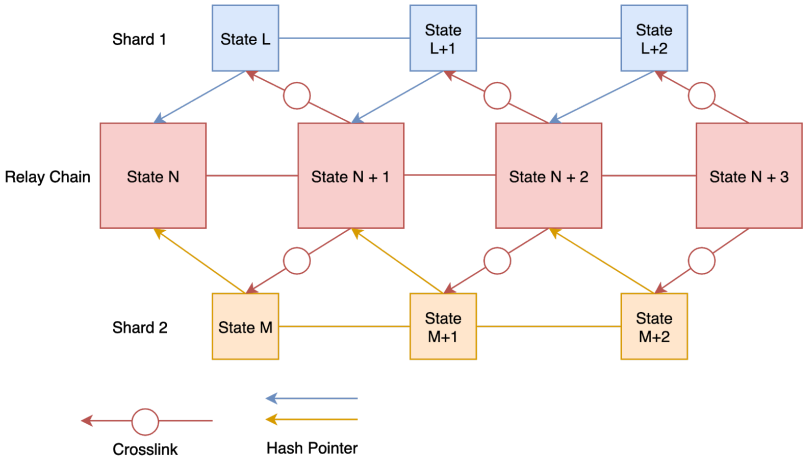


Image 3: A sharded public blockchain

There still exists doubt about whether sharding will work efficiently enough to achieve mainstream adoption, as no sharding scheme has yet to be successfully implemented.[26] However, blockchains do not need in-protocol sharding to rival today's payments rails: layer-two solutions act as free-market versions of sharding, are coming to market now, and will be sufficient to scale blockchains to meet the capabilities of today's payment rails. While layer-two solutions currently exist on top of layer-one chains[27], if sharding is successful, layer-twos will exist on top of shards, which will exist on top of a relay chain.[28]

[25] Stateless clients and erasure coding ensure correct state execution and full data availability, respectively, while secure networking prevents DDoS attacks that could cause system failure.

[26] However, it is likely that sharding will work eventually, as the cryptography and economics necessary for its success is improved everyday by a multitude of global teams, each compounding on the work of others via the nature of open-source.

[27] Layer-two is to layer-one as shards are to relay chains.

[28] This scheme is a form of quadratic sharding.

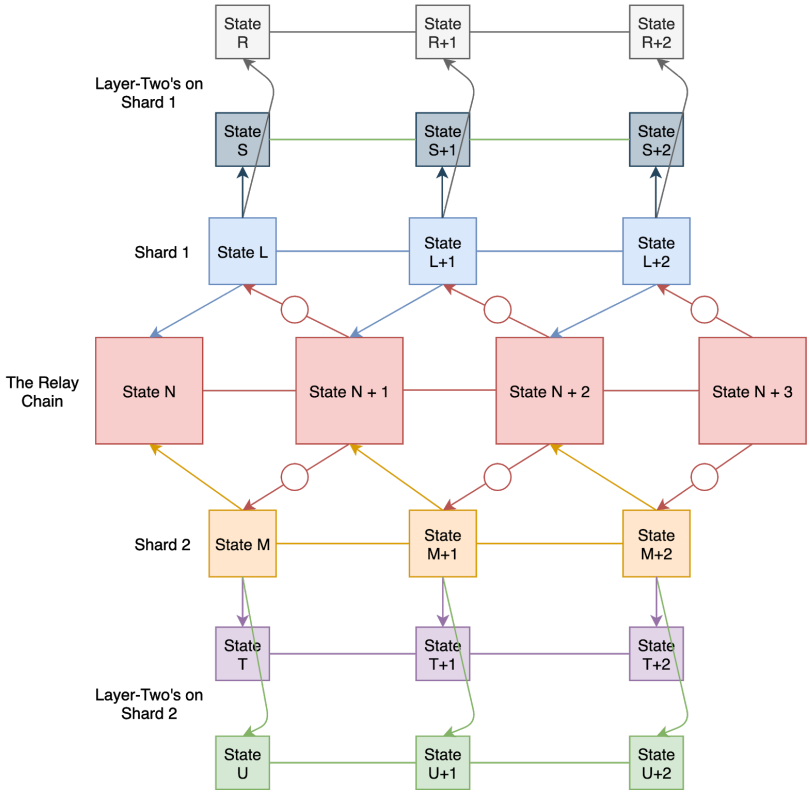


Image 4: A neo-blockchain scaled through a sharded layer-one chain

Privacy

Layer-ones can currently generate ~2 private tx/sec through on-chain mixers. Mixer technology [29] is still nascent, but numerous beta versions are live on public blockchains, one of which is under development by Ernst & Young, one of the largest global accounting firms.[30]

While layer-two solutions have not implemented mixers yet, optimistic rollups promise to do so this year, as optimistic rollups are able to support the same level of computation as their correspondent layer-one chains.[31] Initially, mixers on optimistic rollups may only increase public blockchains' private transaction throughput less than an order of magnitude over today's private throughput of ~2 private tx/sec (We estimate ~10 private tx/sec.). However, optimistic rollups designed specifically for the execution of mixers should eventually be able to support ~200 private tx/sec[32], assuming no enhancement in mixer efficiency.[33]

[29] On-chain mixers use a combination of zero-knowledge proofs and commitment schemes to achieve privacy. Zero-knowledge proofs are the same technology that underpins zk-rollups.

[30] Ernst & Young claims that its mixer can achieve full privacy while only costing $0.05 per transaction. Even though this fee could run significantly higher during times of high network demand (rendering layer-one mixers inadequate for everyday consumer-merchant payments), it would still be two orders of magnitude in improvement over current wire fees for both international businesses and immigrants needing to send remittances.

[31] Optimism's optimistic rollup implementation could do so by the middle of 2020; the layer-two solution is currently in the alpha stage of development.

[32] Modern layer-ones can currently support around 2 private transactions per second. Optimistic rollups should eventually provide ~100 increase on regular Ethereum transaction throughput. Assuming that this same level of augmentation is possible for private transactions via rollups specifically designed for private transactions, we arrive at ~200 private transactions per second.

[33] Optimistic rollups designed for specific use cases can be optimized for performance in ways that generalized optimistic rollups cannot.

However, mixer efficiently will likely continue to improve at a rate of at least 2x per year due to increasing innovation in the zero-knowledge proof space, driving private throughput to within an order of magnitude of today's systems (and with faster clearing/settlement times). Sharding, if successful, would increase the throughput of optimistic-rollup-based private transactions by two more orders of magnitude (over 10,000 private tx/sec[34], assuming no improvement in mixer efficiency).[35]

As the use of mixers proliferates across blockchains, user privacy will be significantly stronger than today's standards: mixer transactions return access control of user data to the user herself; vaults and rails can no longer sell user payment data unless the user specifically gives them access.[36]

Zk rollups will take longer to implement private transactions. In order to achieve privacy, zk rollups must implement recursive zero-knowledge proofs, which, despite the accelerating rate of innovation in the zero-knowledge proof space, is likely multiple years away from being possible. [37] Once recursive zero-knowledge proofs are efficient enough (in terms of proof generation and proof size) to be deployed, zk rollups will be able to scale to the same, if not significantly higher, levels of private throughput of optimistic rollups. Until then, however, zk rollups can provide little to no privacy guarantees at all.

Security

The security of blockchains can be analyzed via two sub-factors: the security of the chain itself (as well as its parts, such as layer-twos) and the security that, even if the chain operates correctly, users are not liable to the consequences of a fraudulent transaction.

Factor One: The Security of the Chain Itself

The security of blockchains rely on both technical security and social security. Technical security is achieved through three disciplines: cryptography, economics, and computer science. Cryptography ensures the validity of all transactions: users send transactions using cryptography; nodes package and verify blocks using cryptography. However, altruism is ineffective in ensuring that every node applies cryptography correctly. Thus, blockchains

[34] 200 private tx/sec * 64 shards = ~10,000 private tx/sec

[35] Initially, shards will be designed to only verify layer-two solutions (not supporting the same level of computation as today's layer-one chains)--this alone could bring throughput to ~10,000 private tx/sec. Eventually, shards seek to support the same level computation as today's layer-one chains eventually, although this will take longer than achieving the ability to only verify the same level of computation possible on today's layer-one chains.

[36] Blockchains can currently trade off enhanced privacy to achieve greater scalability than today's rails (similar transactions per second but with significantly lower fees) through state channels or federated side chains. These solutions are not discussed thoroughly in this paper, as blockchains envision a world without middlemen, which these solutions rely on for their success, similar to VisaNet.

[37] Zero-knowledge proofs are the technology that underpins blockchain-based privacy, both on layer-one and layer-two. One level of zero-knowledge proofs is now possible, as shown through on-chain mixers, but two levels of zero-knowledge proofs has yet to be developed, although numerous teams are researching its implementationA recursive zero-knowledge proof is a zero-knowledge proof that can verify the computation of other zero-knowledge proofs. In the case of zk rollups, instead of the transaction data being posted onto the chain, zero-knowledge proofs that verify those computations could be posted onto the chain, with the recursive zero-knowledge proof verifying that the new state header is the correct result of applying the transactions (represented through zero-knowledge proofs) to the previous state header. State headers are the heads of the Merkle Trees that compress the data of the state of all accounts.

implement economic incentives to ensure that a decentralized community of actors append blocks to the chain that only contain valid transactions: validators receive block rewards for honest behavior and slashing penalties (directly or indirectly[38], depending on the consensus algorithm used) for dishonest behavior. All of a blockchain's cryptographic and economic features are enforced through code, computer science, run by nodes of that blockchain.

Social security entails the ability for honest users to reconvene even in the event of a 'successful' technical attack. In the past (when proof-of-work was the most widespread consensus algorithm), social security relies on shared beliefs, namely that code is not law (as exhibited by the Ethereum hard fork in 2016), and the ability to communicate via online channels (Reddit, Discord, Twitter, etc). Because of social security, blockchains have run securely even if the face of attacks, with honest actors all agreeing to move to the 'honest' chain.

Chains secured by certain consensus algorithms, such as proof-of-stake, can now encode social security (thus morphing it into technical security): honest actors can continue to validate a minority chain by slashing and redistributing that attacker's stake on the minority chain among the honest participants. This significantly increases the security of blockchains, making them almost unattackable by even the wealthiest, most malicious actors.

Other consensus algorithms, such as proof-of-work, do not provide the same ability to technically reconvene on the minority chain[39], and so they rely on greater community participation on online forums to arrive socially on the correct chain. This is less secure, as un-encoded social security can lead to significantly more variation than encoded social security, hence why proof-of-stake appears to be the dominant consensus algorithm moving forward.

As for layer-two solutions, optimistic rollups act as another blockchain on top of the layer-one blockchain and thus follow similar properties (although there are some differences, specifically with regards to how to merge technical and social security), while zk rollups are an exception to using economics for security.

Zk Rollups

Zk rollups garner their security properties from validity proofs, which leverage cryptography alone to prove correctness. This feature prevents zk rollup validators from posting invalid transactions, which in turn allows for users to exit from a zk rollup back to its correspondent layer-one in the layer-one's next block (~13 seconds in Ethereum's case). Some zk rollup teams have implemented economic incentives to enhance zk rollups' scalability. Economic incentives in zk rollups allow for sub-second, guaranteed transactions (at the scale of 2000 tx/second). However, there is a strong difference between enhancement via incentives to perform optimally and a reliance on incentives to perform securely, as in the case of optimistic rollups.

[38] Either through a direct loss of capital (proof of stake) or an indirect loss of capital through wasted electricity consumption (proof of work) or wasted storage space (proof of storage).

[39] A miner can continue to attack the mintory chain unless the minority chain changes the hash algorithm that underpins their proof-of-work algorithm.

Optimistic Rollups

Optimistic rollups rely on two blockchains for security: the rollup and its corresponding layer-one. To provide adequate security assurances through economic incentives, optimistic rollups implement challenge periods (via their smart contracts on layer-ones), where validators can challenge an optimistic block by locking funds in a smart contract. Challenges reward honest actors and punish bad actors (through distributing either the block proposer's stake or the challenger's stake to the honest party, similarly to layer-one proof-of-stake chains 'in-protocol').

The implementation of challenge periods increases user exit times from optimistic rollups back to layer-one chains significantly: several hours (lower bound) to one week (upper bound). Sufficient challenge period length is extremely important for security of optimistic rollups. Proof-of-stake optimistic rollups cannot rely on the same social securities as proof-of-stake layer-ones, as, if an attacker successfully compromises and exits an optimistic rollup, the layer-one will not recognize it as an attack. Long challenge periods mitigate this risk, as they allow honest validators to have enough time to find and confirm dishonest behavior.

Proponents of optimistic rollups have developed workarounds for their suboptimal exit times: to establish a credit market in which credit providers settle user exits to layer-ones for a fee. This credit market would have little implementation difficulty due to the advent of submarine swaps[40] and the creditor's ability to both monitor and verify the optimistic rollup's state in real time, allowing the creditor to be confident that borrower's layer-two funds would become their own should the creditor execute the swap. Credit markets decrease exit times of optimistic rollups from hours-days to minutes.

Thus, with the introduction of additional economic incentives on top of layer-one's incentives, optimistic rollups can achieve scalability, privacy, and security in the near future. Consequently, optimistic rollups appear to be a better near-term solution for a disruptive blockchain-based payment system, while zk rollups appear to be a better long-term solution due to their decrease in dependencies on economic incentives, relying neither on fraud proofs or credit markets. However, in the defense of optimistic rollups, a reliance on economic incentives has not yet been a problem for layer-one chains.

	Layer-One: a public blockchain	Layer-Two: Zk Rollups	Layer-Two: Optimistic Rollups
Scalability	Can currently support tens of transactions per second, each cleared in 13 seconds and settled in two minutes, charging cents to dollars in fees.	Can currently support thousands of transactions per second, each cleared in a second and settled in 13 seconds, charging sub-cent fees.	Can currently support upwards of 500 transactions per second, each cleared in 13 seconds and settled in two-four minutes, charging sub-cent fees.

[40] Submarine swaps allow for trustless exchange of the creditor's layer-one money for the borrower's layer-two money.

Privacy	Can currently provide privacy through on-chain private mixers (~2 tx/sec).	Requires significant innovation in zero-knowledge proofs, which may be multiple years away.	Can provide privacy through layer-two mixers this year (~10 tx/sec). Should scale by at least an of magnitude over the next two years.
Security	Extremely high to due encoded technical and social security.	Relies on validity proofs. Maintains the trust assumption to only the layer-one validators.	Relies on fraud proofs and credit markets. Increases the trust assumptions to the layer-one chain *and* layer-two validators.

Factor Two: Anti-Fraud Solutions

The other significant aspect of security remains the same regardless of which layer-one or layer-two solution is used to transact: public blockchain infrastructure must combat fraudulent payments.

Current payment networks often place the risk of fraud onto the merchant, reversing payments to a merchant who may have had no idea they were being frauded. From first principles, placing the risk of fraudulent payment onto the merchant, who already loses margins because of fees paid to payment networks, is a user behavior that only exists because of the monopolies afforded to current payment networks.

Payments cleared on blockchains are irreversible. Thus, blockchains allow for a more rational solution: the risk of fraudulent payments is passed to the party that authorized the payment, either the user themself or the user's digital wallet, should the user enlist in the wallet's insurance solution. Digital wallets may provide this insurance for 'free'.[41]

Users will have the ability to choose between the two. Those that choose to authorize their payments alone (hopefully using some form of two-factor authentication) will likely be technologically savvy users, while mainstream users will choose to work with digital wallets.[42]

Security of Centralized versus Decentralized Systems

Any system, whether centralized or decentralized, is composed of multiple parts, and each system type is only as secure as its least secure part. Decentralized systems inherently account for this in their designs. They build resilience into each part, with each part treating every other part as adversarial. Thus, if a hacker successfully attacks a part of a decentralized system, the

[41] Bio-authentication of blockchain-based payments and sufficient anti-fraud algorithms (based on location and past purchasing habits) may be enough to drive the probability of fraud extremely low.

[42] Should the provider detect that the transaction has a high probability of fraudulence, the provider would notify the user and not make the payment until receiving the user's approval. Should the insurance provider approve a transaction signed by someone pretending to be the user, the insurance provider would be liable to pay the user. However, it is extremely difficult to pretend to be the user in a blockchain-based world, as the malicious actor would need access to the user's private key, which should either be stored in a safe or embedded into hardware (such as a smartphone), only accessible through biometric verification.

hacker can only compromise that singular piece of the system. As decentralized systems grow in their number of parts, the security of each part remains the same.

However, centralized systems only protect their perimeters: internal parts are allowed to communicate with each other without assuming that their corresponding part could be adversarial. As centralized systems grow in their number of parts, the size of their perimeters increases as well, and consequently so does the attack surface for a hacker. Thus, as centralized systems grow, the security of each part, as well as the system more broadly, decreases. Consequently, blockchains (properly designed decentralized systems, in general) have significantly stronger security guarantees than centralized systems.

Triggering the S-Curve

Blockchains will eventually deliver more scalable[43], more private[44], and more secure[45] payments than today's payment infrastructure, but it will take multiple years before these three features can be optimized simultaneously, as teams continue to innovate on horizontal scaling solutions and zero-knowledge proof efficiency.

There are numerous actors that stand to benefit from blockchain payments as they exist today and in the very near future: international businesses, migrant workers, and citizens living under oppressive regimes. The first two parties stand to achieve significant efficiency gains, both due to cheaper fees and faster settlement times.[46] The last group may be willing to put up with blockchains' present shortcomings if their alternative is losing a significant amount of wealth from inflation.

The Future of the APIs

As blockchain-based banking/payment solutions proliferate, the APIs of today's financial stack will be rendered irrelevant. The open-source ethos dominating public blockchain infrastructure renders obsolete both categories of current fintech APIs (rail APIs and data APIs).

Rail APIs, such as Square's and Stripe's, are made insignificant due to the open-source nature of layer-one and layer-two solution payment solutions: the code necessary to use these networks is already published, free of charge.

---

[43] Layer-one's scalability is being enhanced by the widespread transition to proof of stake consensus algorithms as well as implementation sharding, while layer-two's scalability is being enhanced by innovation in zero-knowledge proofs and virtual machine optimization.

[44] Both layer-one and layer-two's privacy are being enhanced by continued innovation in mixers and the zero-knowledge proofs that underpin them. Blockchains, in this case, are more private because they return access control of user data to the user herself. Visa can no longer sell user payment data unless the user specifically gives Visa access.

[45] Decentralized systems are inherently more secure than centralized ones.

[46] It is reasonable to assume that corporations have an average cost of capital of 15% (while banks, hold the funds as they are being transmitted, only have an average cost of capital of 3%). This 12% discrepancy equates to 3 basis points (0.03%) per day of inefficient capital. When compounded daily, 3 basis points becomes 6 basis points on day two (the average time for an international transaction) and continues to rise by ~3 basis points every day the transaction has yet to settle thereafter. If a business sends at least one $100k wire per day, that business bears a cost of capital of at least $60 per wire (generously assuming a two day settlement time). Combined with the savings from avoiding upfront wire fees, blockchain-based payments can save large, multi-national organizations hundreds, if not thousands, of dollars daily (This amounts to tens to hundreds of thousands annually.). Additionally, the transparency of blockchain eliminates unknowns such as the timing of settlement (currently, organizations have little to no insight on when their transaction will settle), which will allow firms to better schedule their business activity.

Data APIs, such as Plaid's, exploit a lack of interbank standards with regards to data access control. Blockchains change the data ownership structure entirely: users control access to their data, instead of commercial banks. As to how digital wallets access consumers' blockchain-based data will depend on the privacy solution (or lack thereof) used by the consumer:

- If the user does not utilize any privacy solution, all of that user's transaction history would be public via his or her public address(es) on a public blockchain. Thus, a digital wallet could query the entire blockchain for the user's activity through either a free-market query protocol, such as the Graph or Infura, or its own internal system.
- If the user utilizes zero-knowledge proof based privacy solutions (such as on-chain mixers or a recursive zero-knowledge proofs), the user could store their encrypted transaction history[47] on a decentralized storage protocol, such as Filecoin, and authorize access to digital wallets through a proxy re-encryption protocol, such as NuCypher.[48]

These protocols (query, storage, proxy re-encryption, as well as any other blockchain-based protocol) allow for free markets of services, consequently preventing any party from monopolizing access: anyone who raises fees excessively will be undercut by a competitor with little friction, as the very nature of a protocol allows for near-frictionless interoperability between service providers.

It will be up to consumers to choose how they manage their financial data, but all of the back-end interaction between query protocols, re-encryption protocols, blockchains, and decentralized file systems will be completely abstracted from them by digital wallets. In the near-term, our new financial stack is the following ('no arrows' indicates a double arrow):

[47] The user's transaction history would be encrypted by her viewing key, which the user could either store themselves or entrust to a centralized party (such as a digital wallet) or a distributed key generation protocol.

[48] Proxy re-encryption protocols are an innovation in access control, using novel public/private key schemes that allow only the intended party to view the decrypted version of encrypted data, without any centralized middleman necessary.

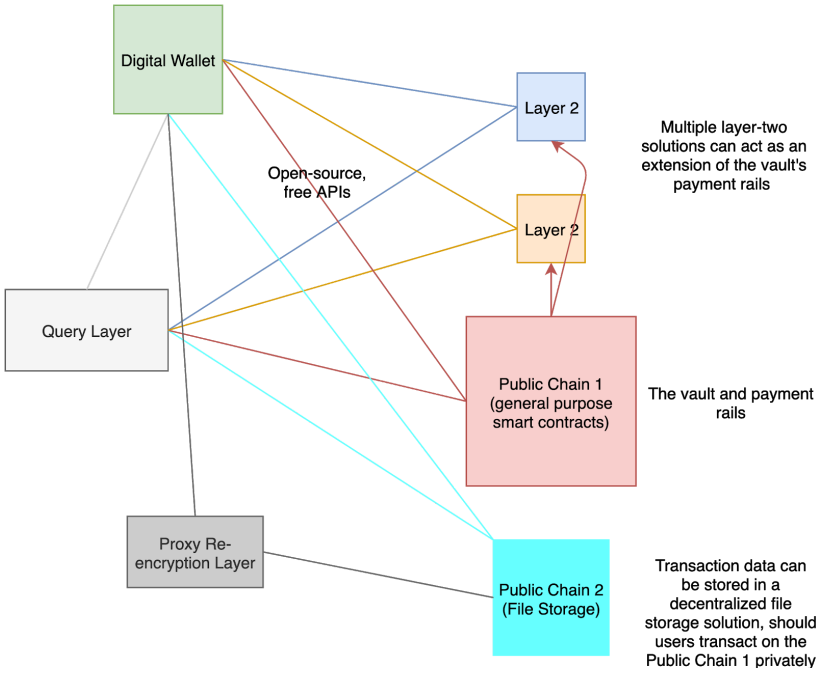


Image 5: Future fintech infrastructure without sharding

If sharding is successful, our new financial stack becomes the following:

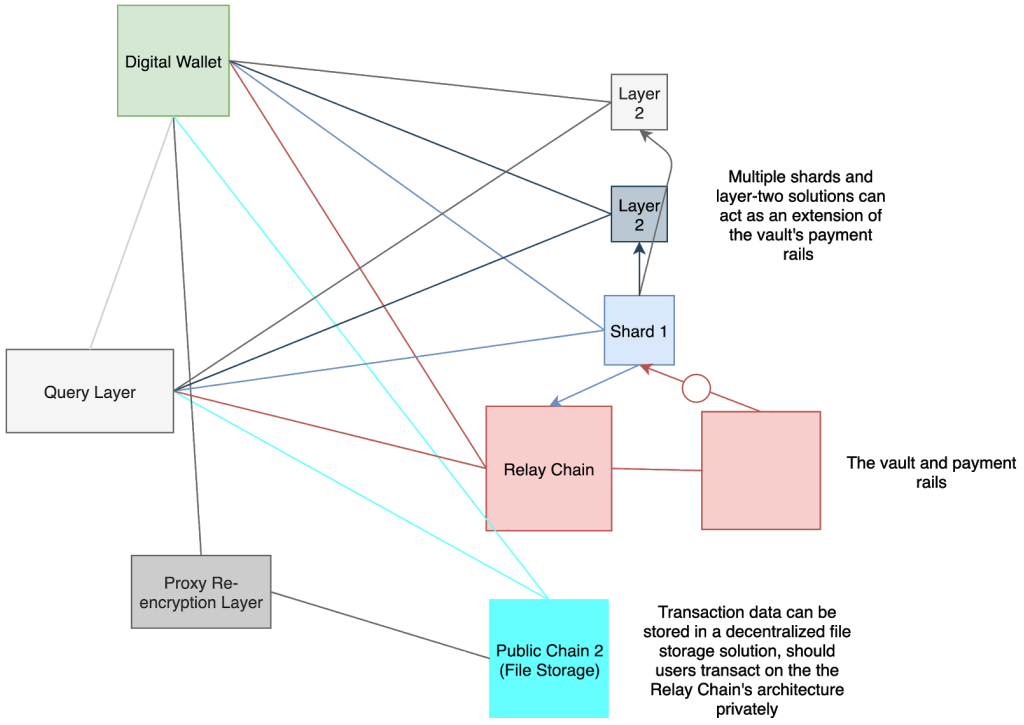


Image 6: Future fintech infrastructure with sharding

Conclusion

There are three layers of fintech--the vault, the rails, and the APIs. Each of these layers will be disrupted by decentralized finance. While each of these three layers possess numerous moats, history serves as a reminder that all moats are temporary. For example, telecommunication companies also benefited from significant regulatory and network moats, and yet they could not do anything to prevent the advent of the internet. The largest open question that remains is whether layer-one chains will interface with traditional vaults (through stablecoin issuers' DBSs), federal governments' proprietary chains (through central banks' CBDCs), or remain fully sovereign (through a decentralized communities' FCOS). Blockchains leave little room for the incumbents of the payment and API layers, other than their becoming providers on top blockchain-based protocols (such as a validator of layer-two rollup or a 're-encryptor' in a proxy re-encryption protocol).

However, the power of incumbents will not disappear suddenly or completely. There always exists a strong faction of users that are slow to adopt new technology, and incumbents will continue to serve these people. That said, efficiency wins overtime, as people will learn new behaviors to save money and time. Thus, risk management practices would suggest that these incumbents should bet on public blockchain infrastructure, in order to de-risk their overexposure to outdated technologies. The only way for them to do this, unfortunately for them, is to buy the assets that secure these public blockchains. Other than the assets that secure public blockchain infrastructure, digital wallets stand the most to gain.

Fintech is already converging inside digital wallets, such that payments, savings, budgeting, investments[49], and other financial services are all accessible in application. Digital wallets will benefit from the free-market economics that public blockchains provide below them: reduced fees at the infrastructure layers will allow digital wallets to produce their services at incredibly low costs. Additionally, digital wallets will be empowered to conduct further value-add services, such as abstracting away all of the complicated user behavior required to make blockchain-based finance work: custody solutions, anti-fraud solutions, and the managing of access to users' private data. However, these front-ends will not benefit from moats in the same way that current banks have, as the barrier to enter the digital wallet space will be significantly lower than it is today, another advent of public blockchain infrastructure.

Special thanks to March Zheng, Michael Stalder, Gary Thung, Maximillian Jungreis, and Ankit Goyal for their feedback.

[49] Popular investment instruments (such as equities, debt, or real estate), which do not currently exist en masse on blockchains, will eventually be settled on blockchains: users will be able to atomically exchange stablecoins on one chain for securities issued on either the same or another chain.