Scaling Blockchains

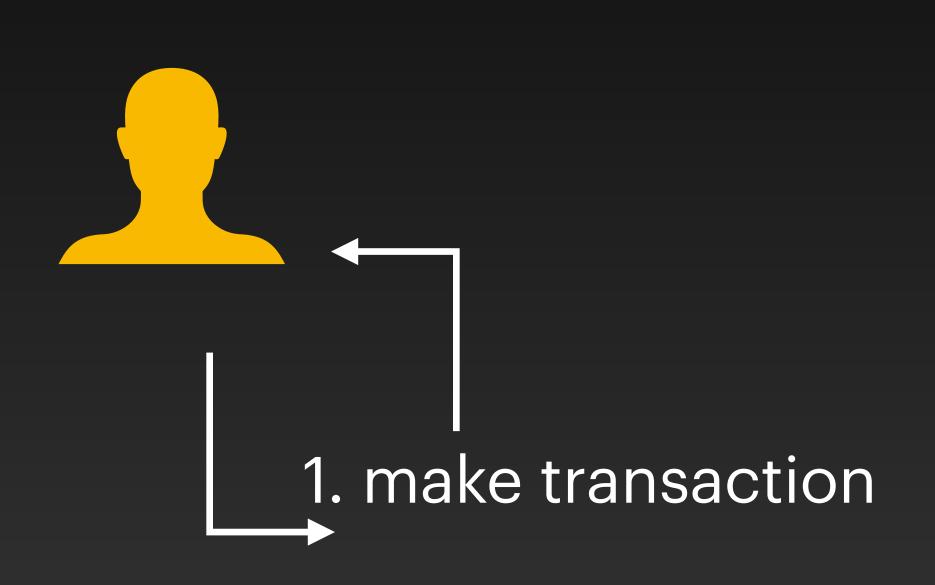
Achieving Arbitrary Throughput and Sub-Second Latency

A set of nodes

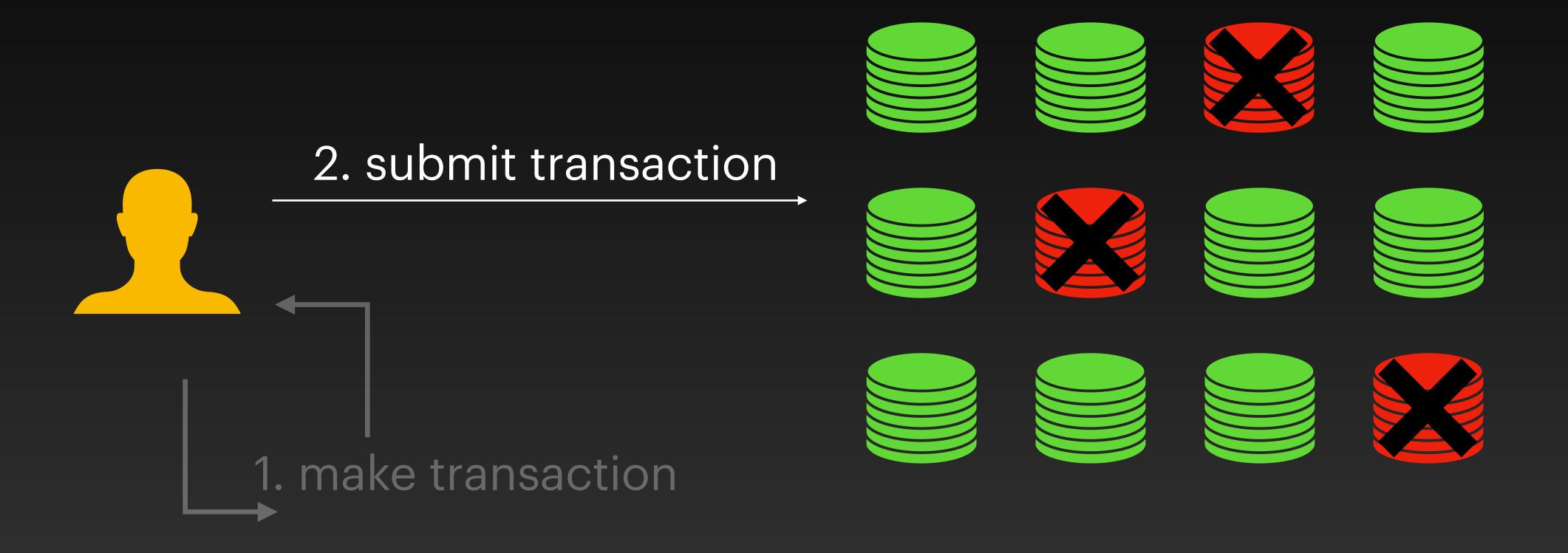


Byzantine Fault Tolerance

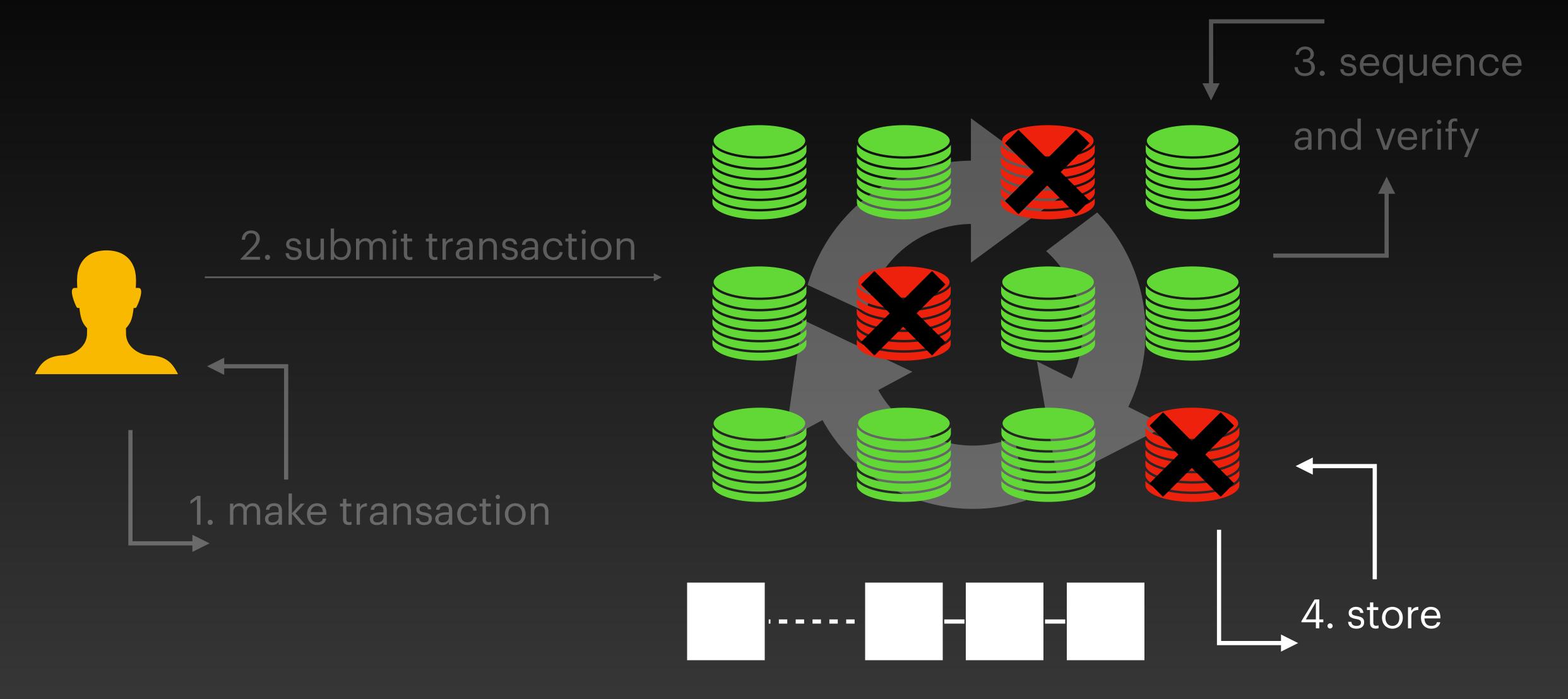




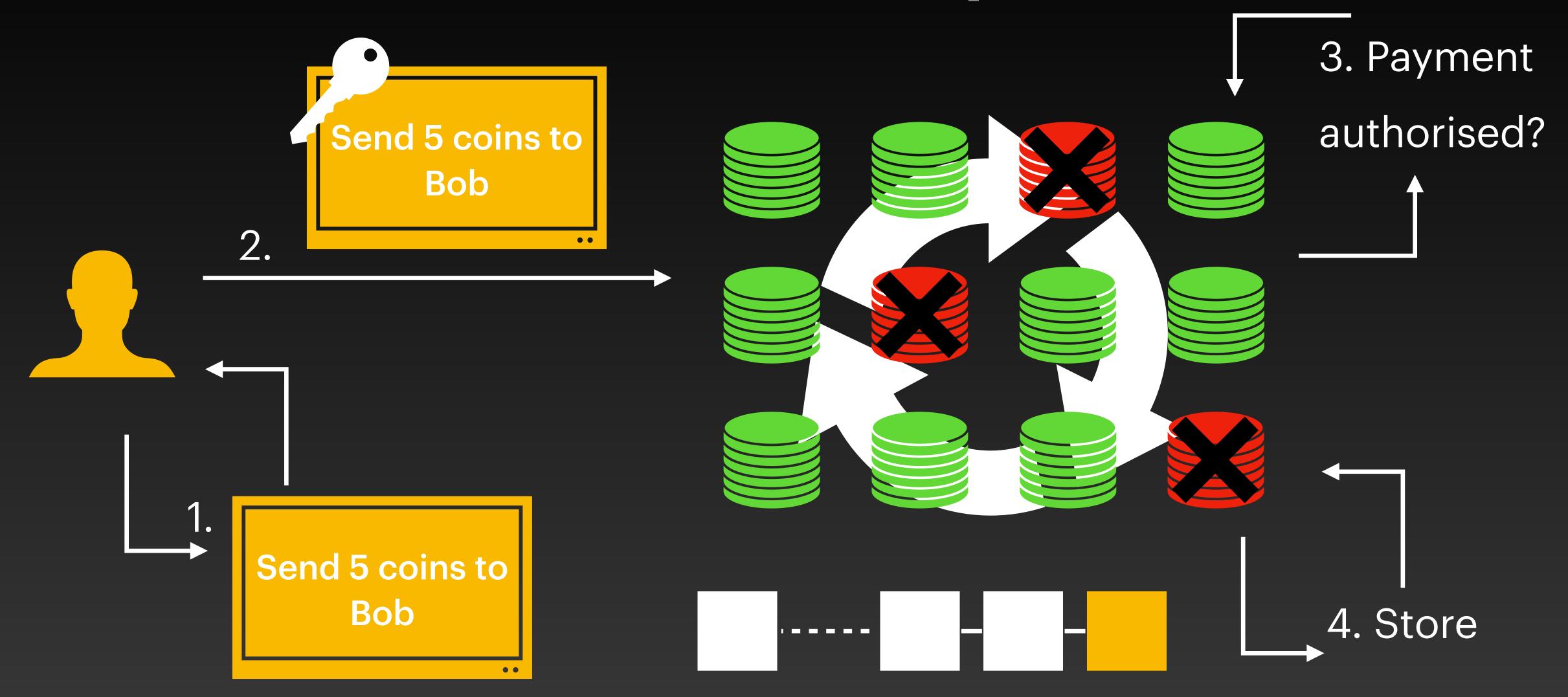








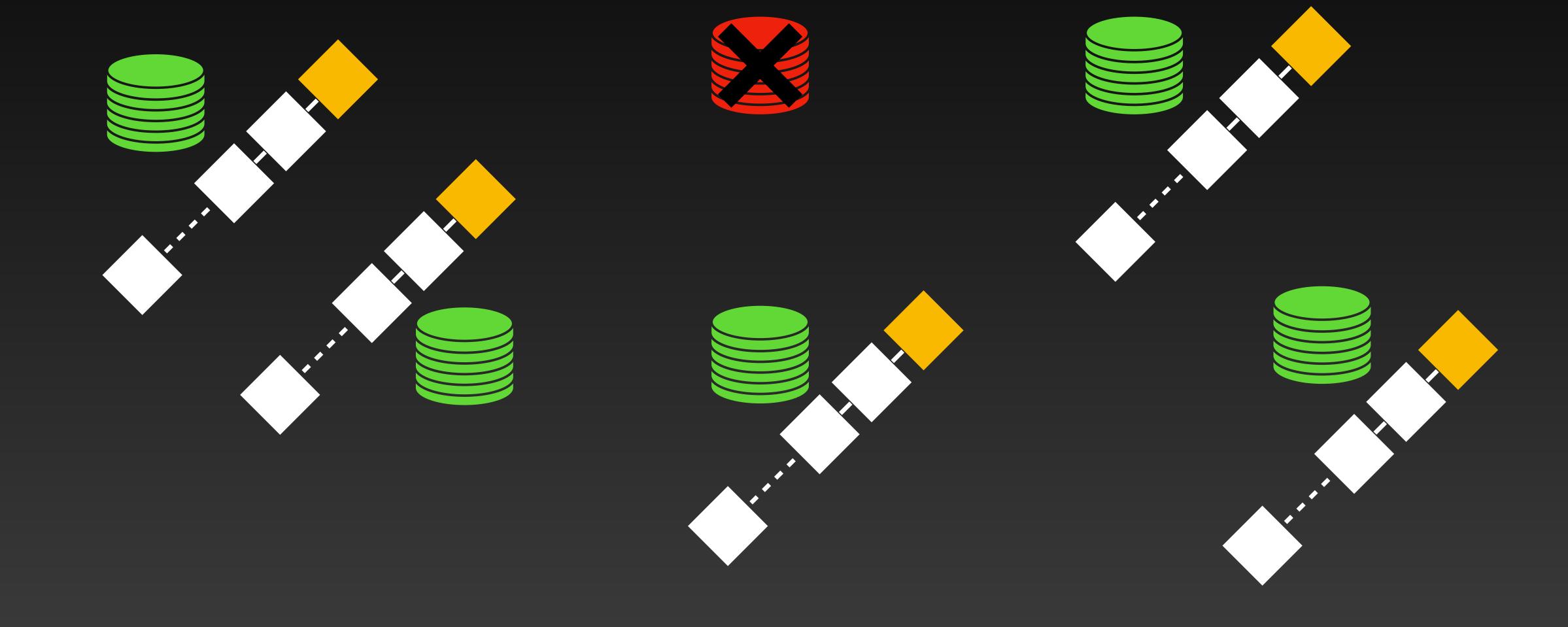
The best example



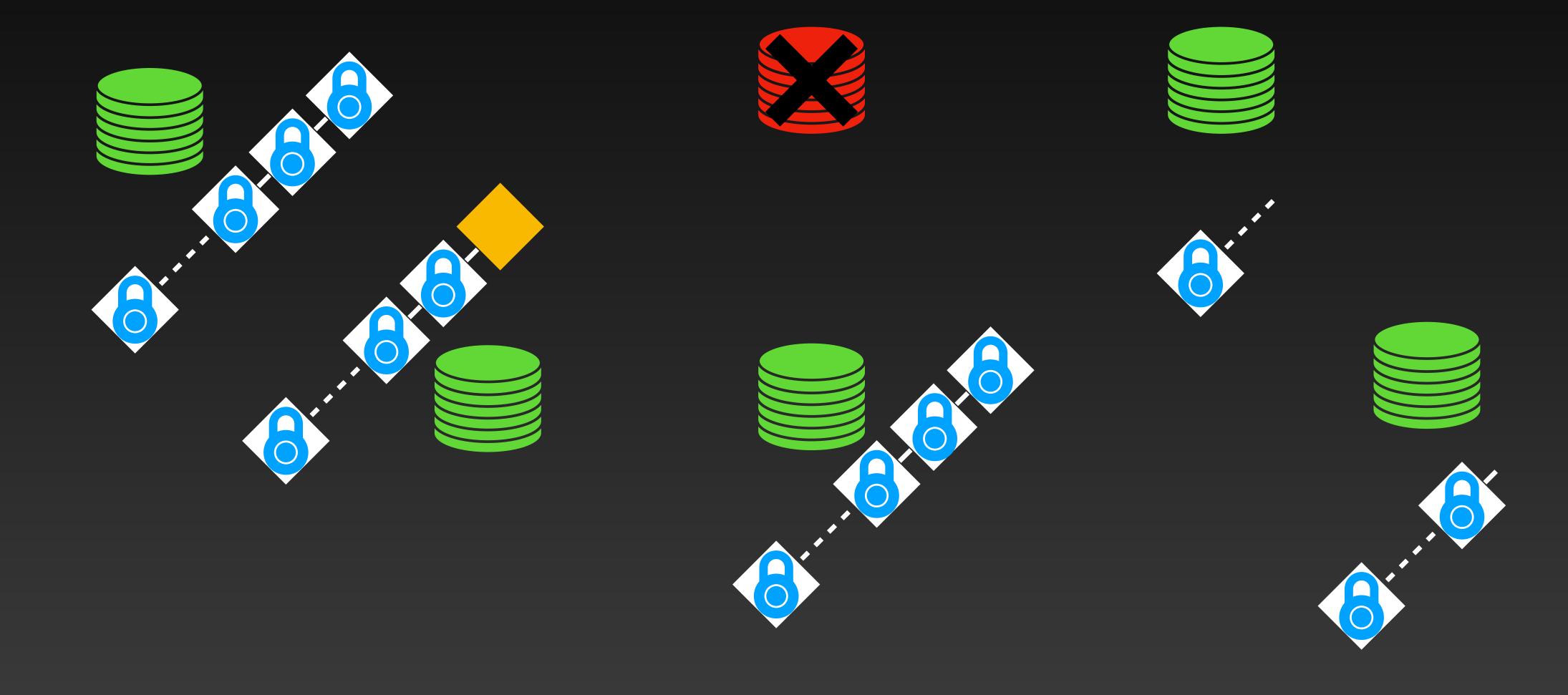
High latency
Slow finality

Low throughput

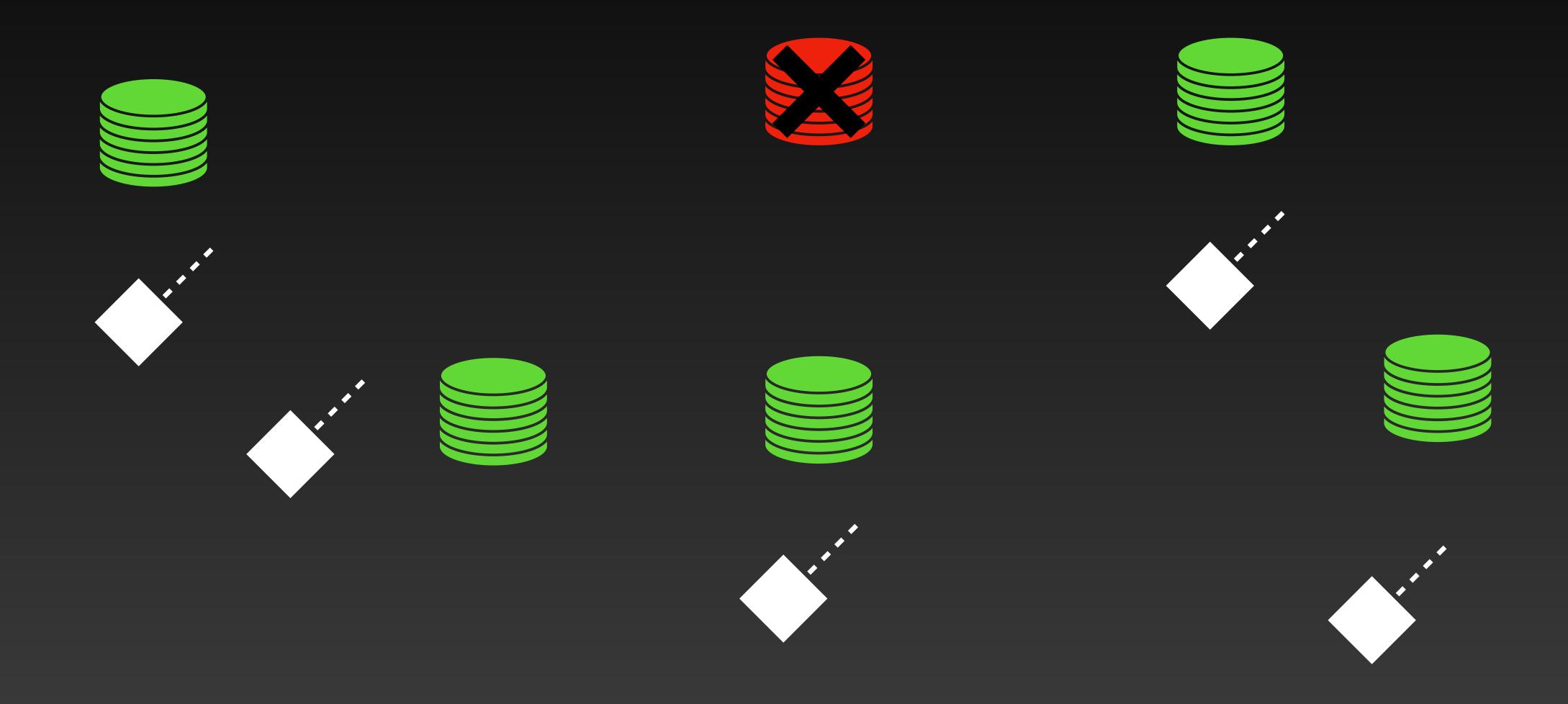
Hight Latency



Slow Finality



Low Throughput



Outline

Part I: Increase throughput

Part II:
Reduce latency

through sharding

with side infrastructures

Outline

Part I: Increase throughput

Key concepts of state sharding

Cross-shard consensus

Replay attacks against common systems

Secure Cross-shard consensus

Evaluation

Part II: Reduce latency

Outline

Part I: Increase throughput Part II: Reduce latency

Blockchains for retails payments

FastPay as side infrastructure

Interfacing FastPay with a primary system

Implementation

Evaluation

Main Takeaways

Part I: Increase throughput

Key concepts of sharded distributed ledgers

Main challenges in building secure sharded systems in practice

Part II: Reduce latency

Side-infrastructures to bring blockchain-based payment systems to physical points of sales

How to integrate those infrastructures into a primary distributed ledger

What is not covered

Privacy on blockchain

Sybil resistance mechanisms

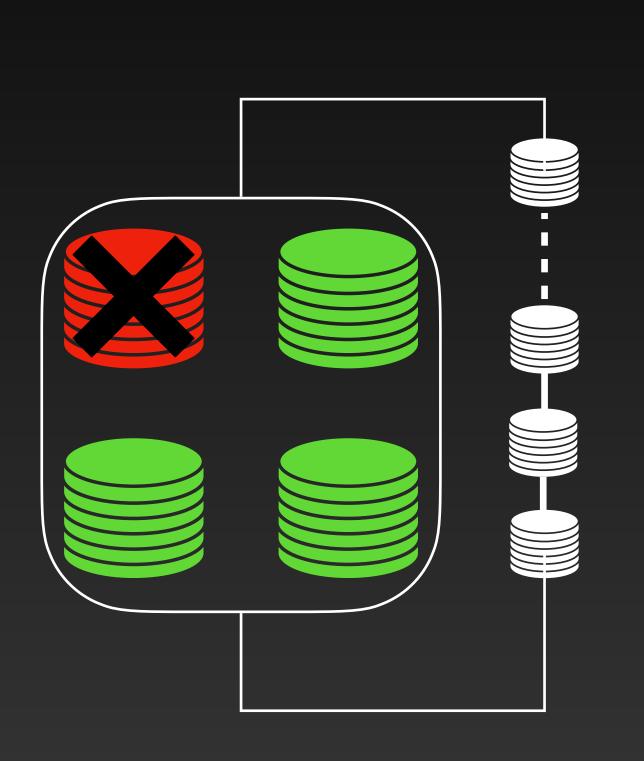
Incentives of nodes operators

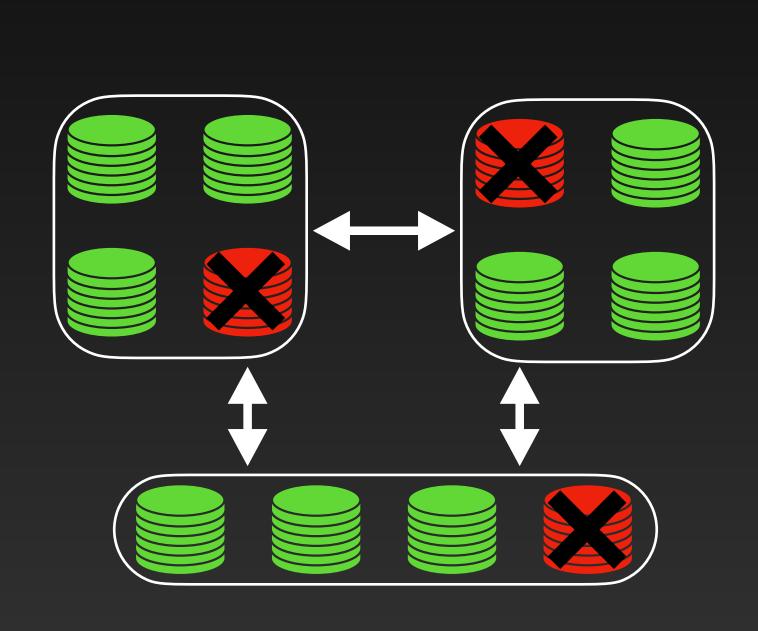
Increasing Throughput

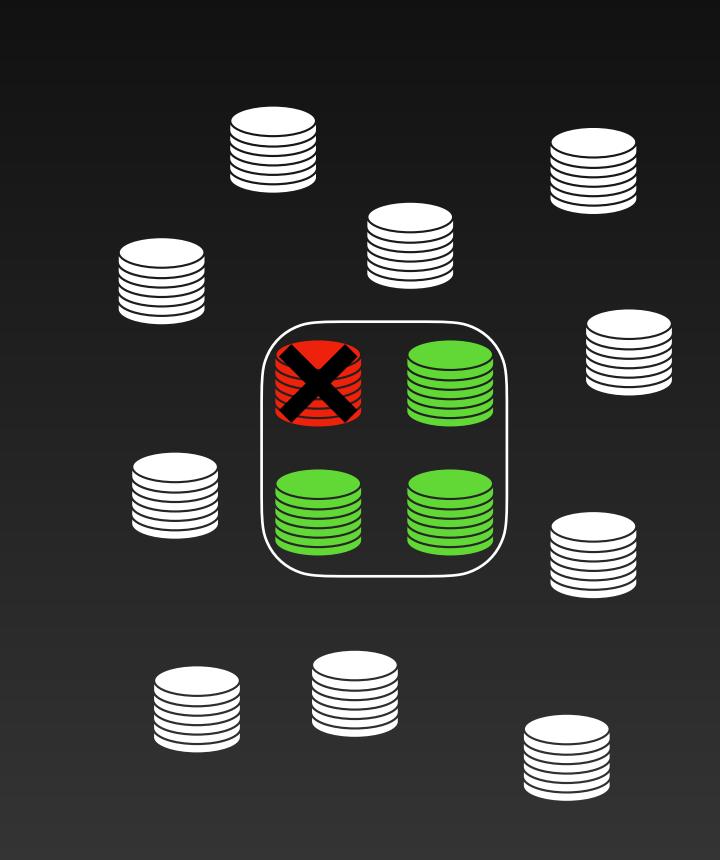
with Sharded Blockchains



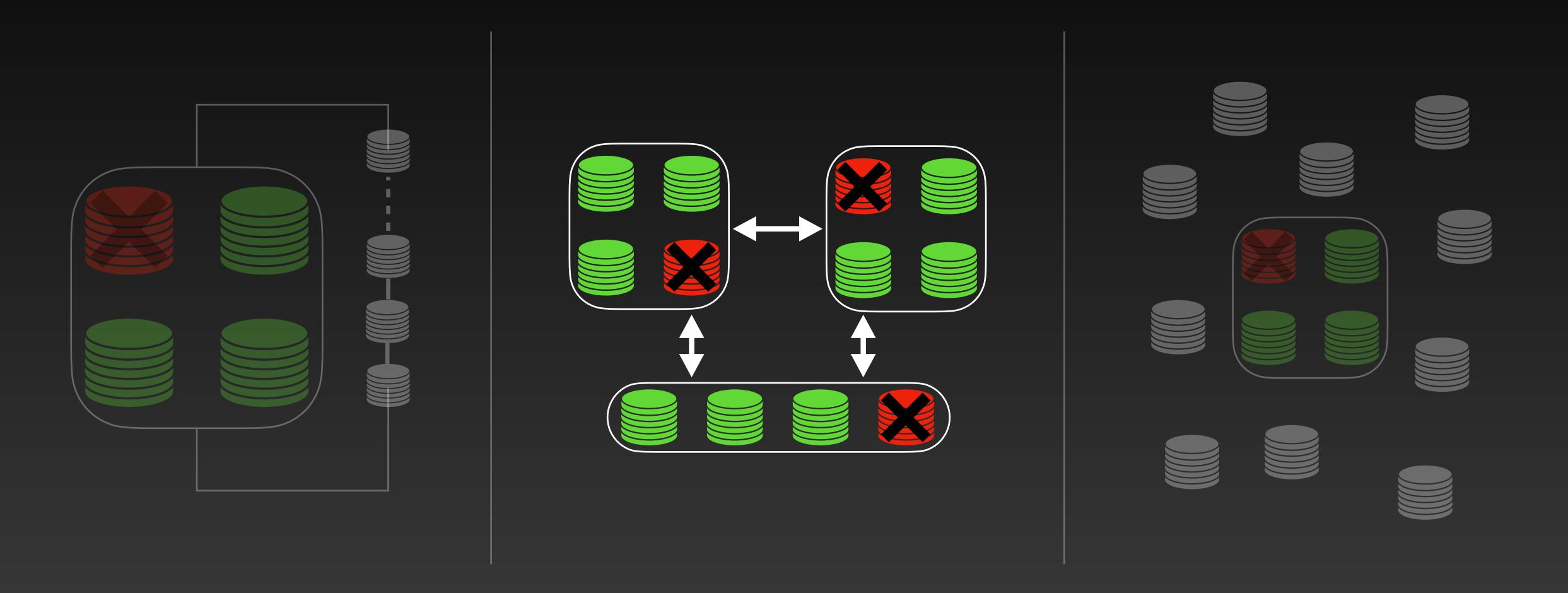
Scaling blockchains







Scaling blockchains



Hight throughput

BFT resilience

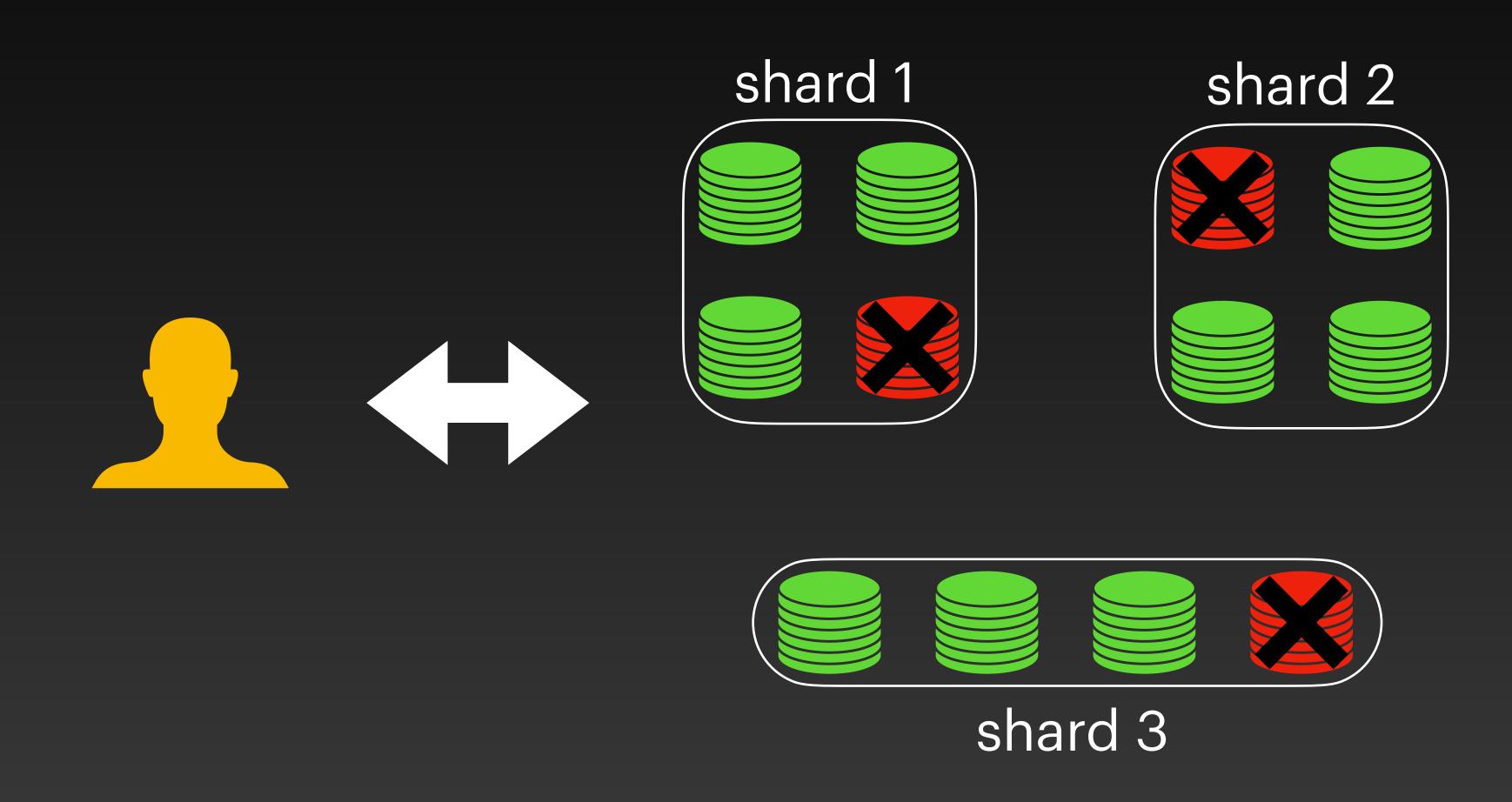
Fast finality

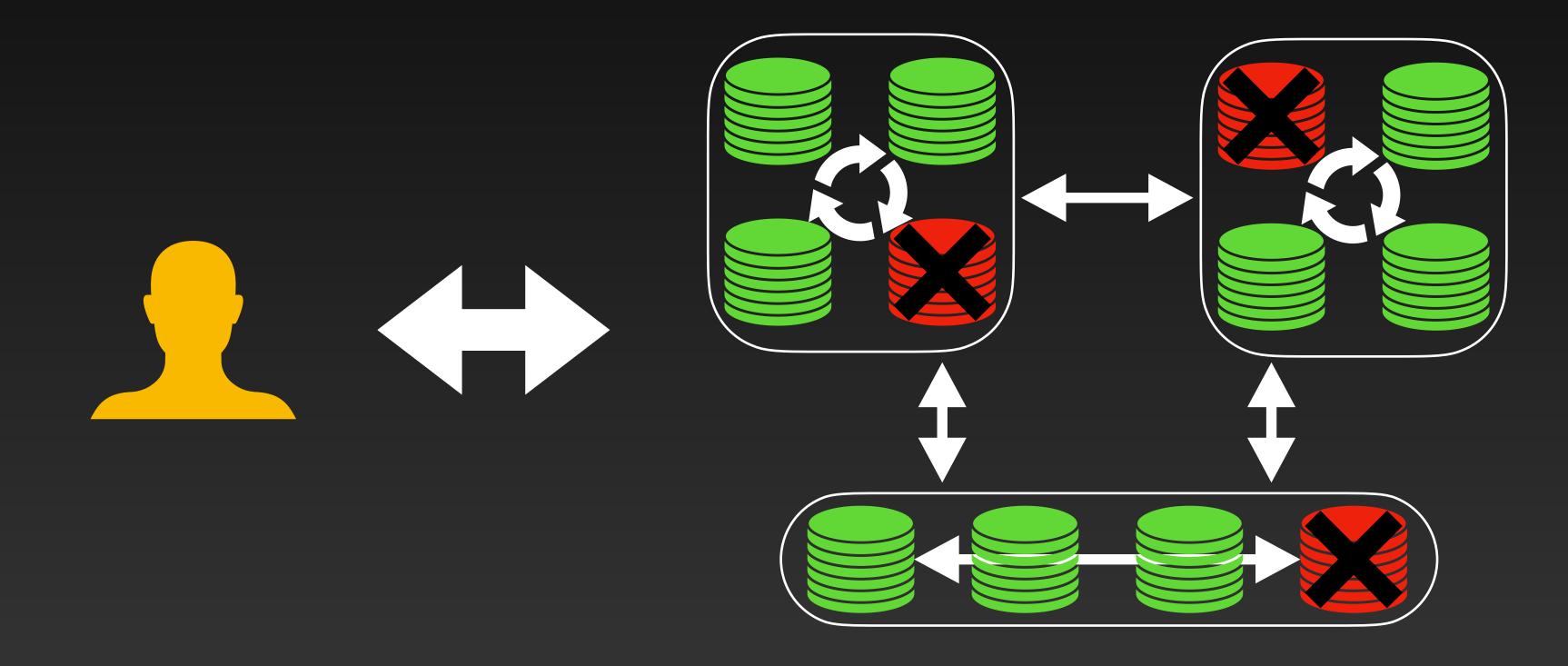
Linear scalability

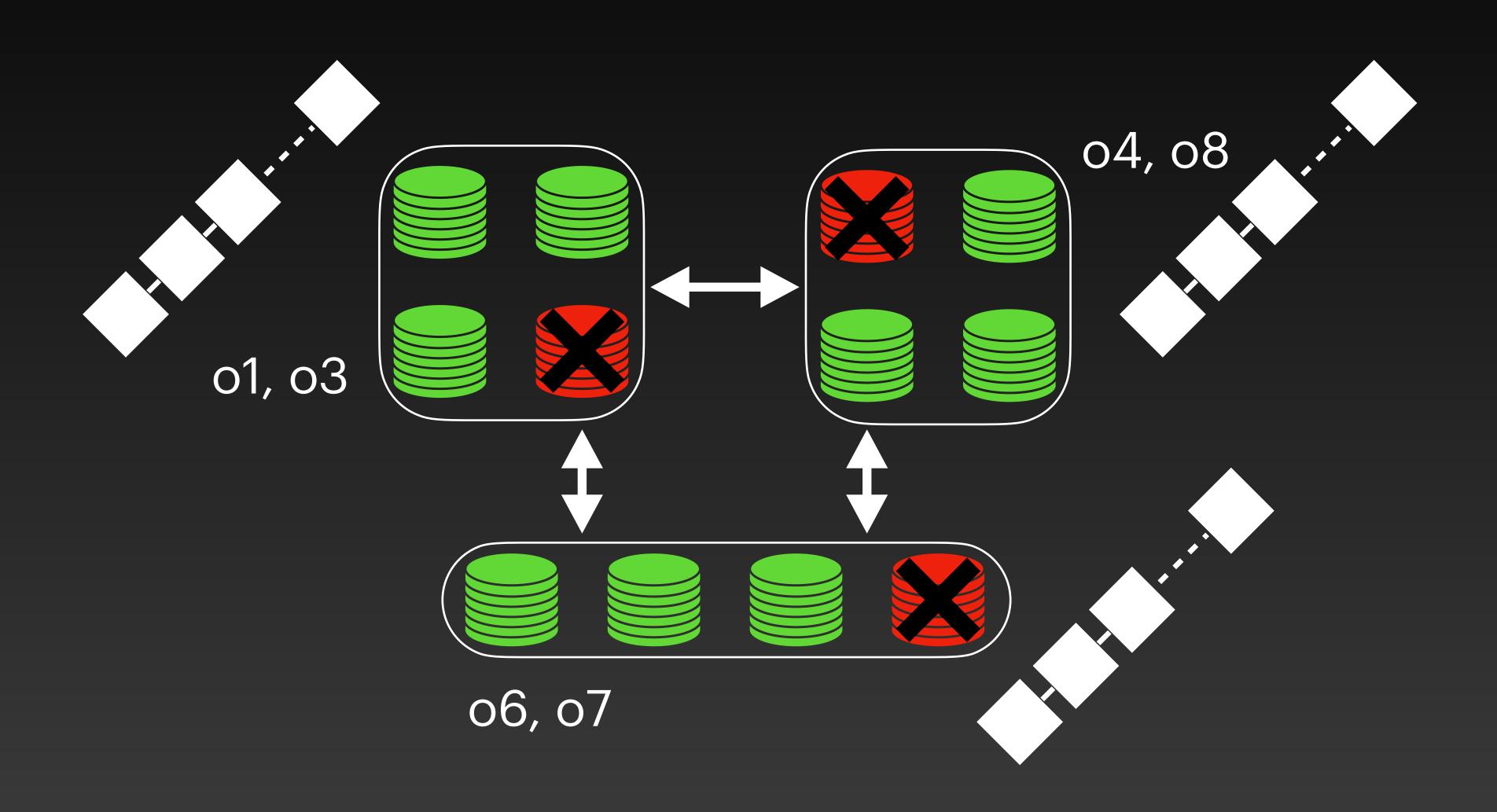
The more machines you have, the bigger your throughput

Scalability







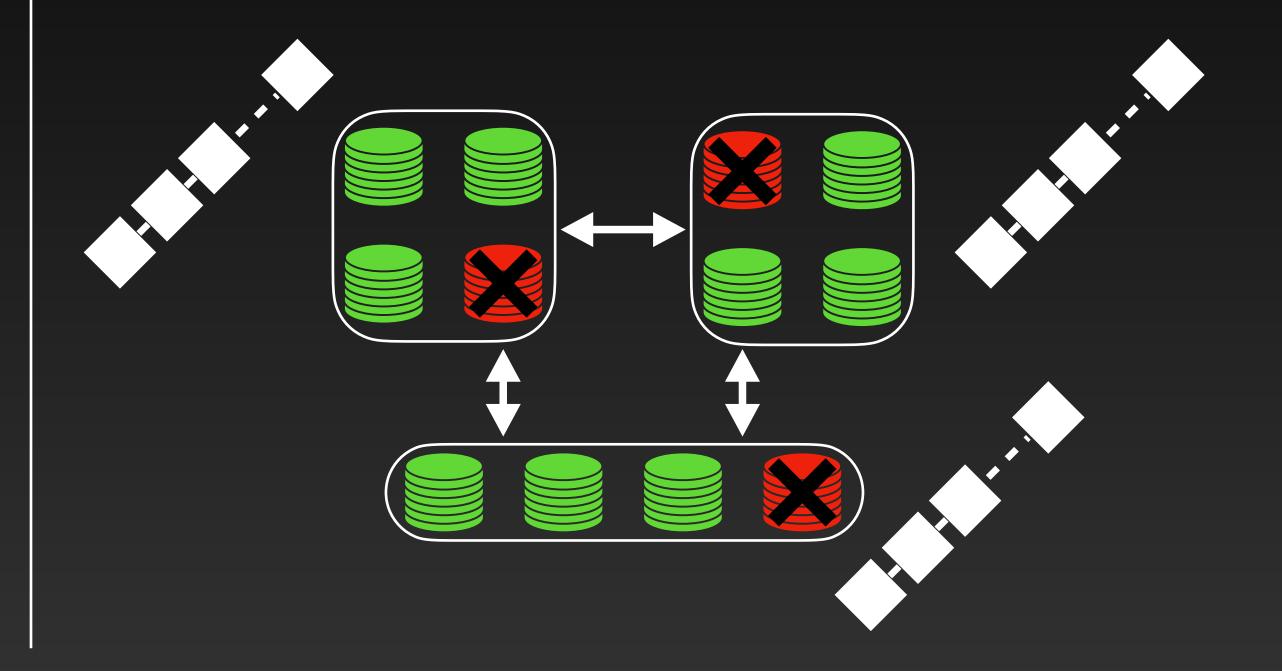


Traditional



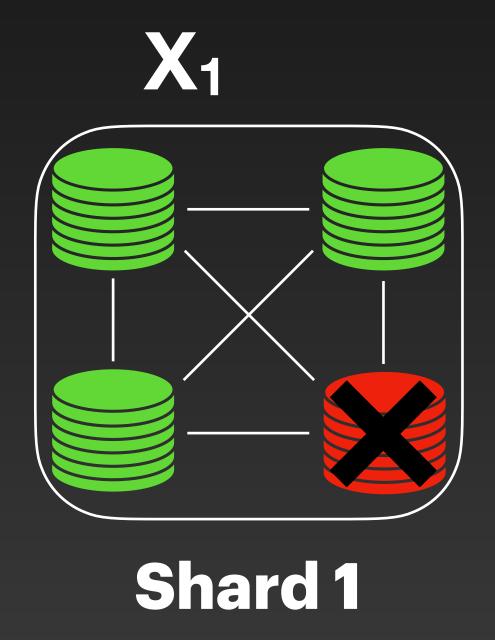


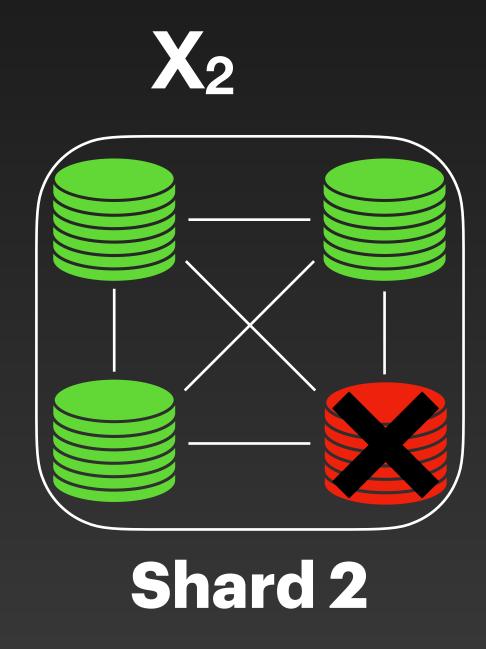
Sharding



An example transaction

$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$

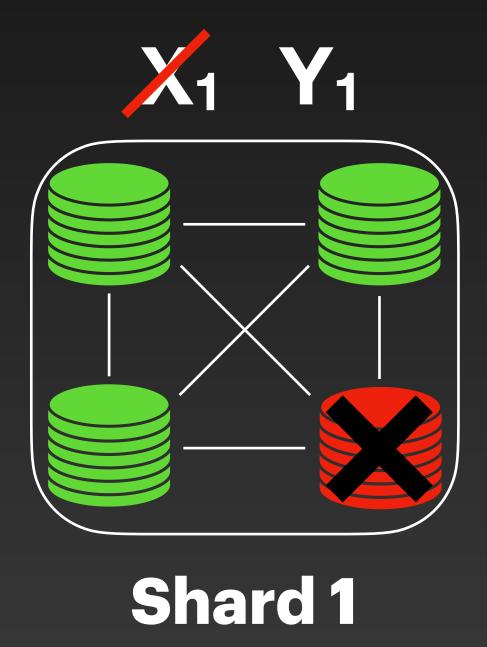


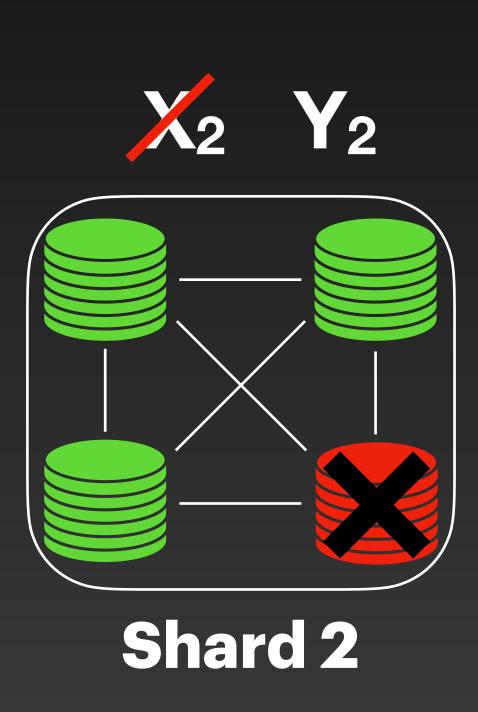




An example transaction

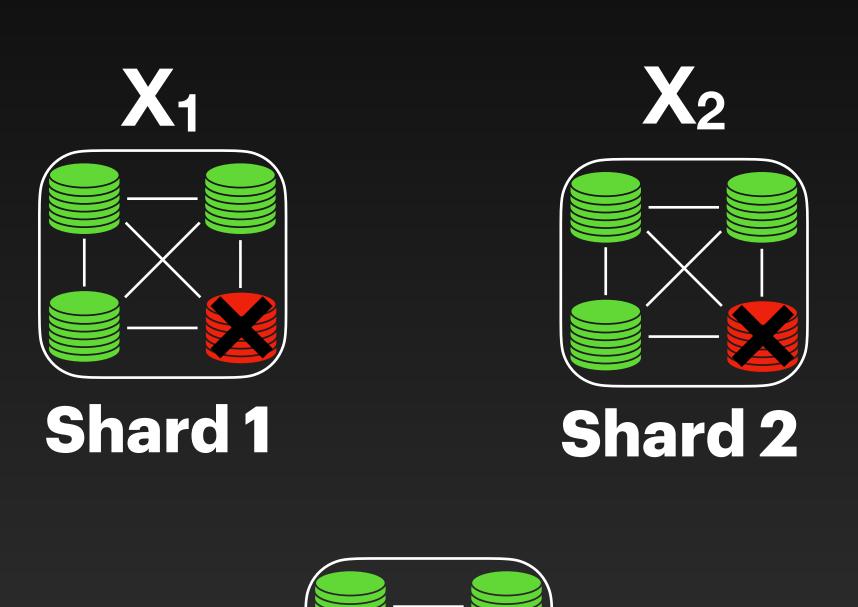
$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$



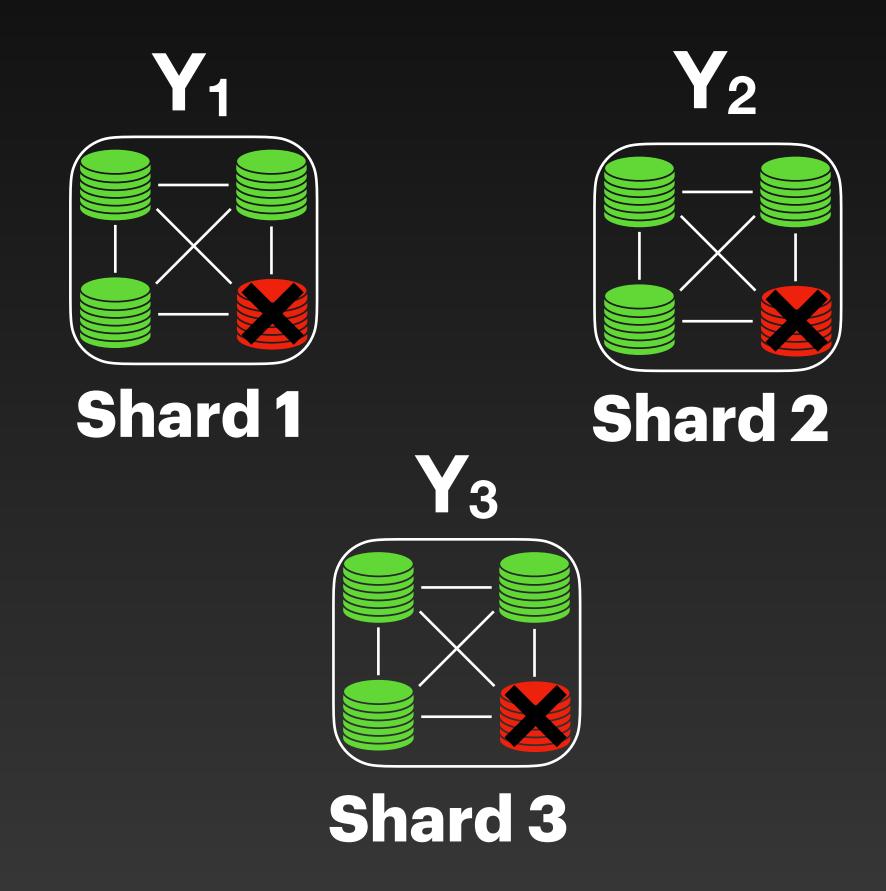




Only two acceptable final states

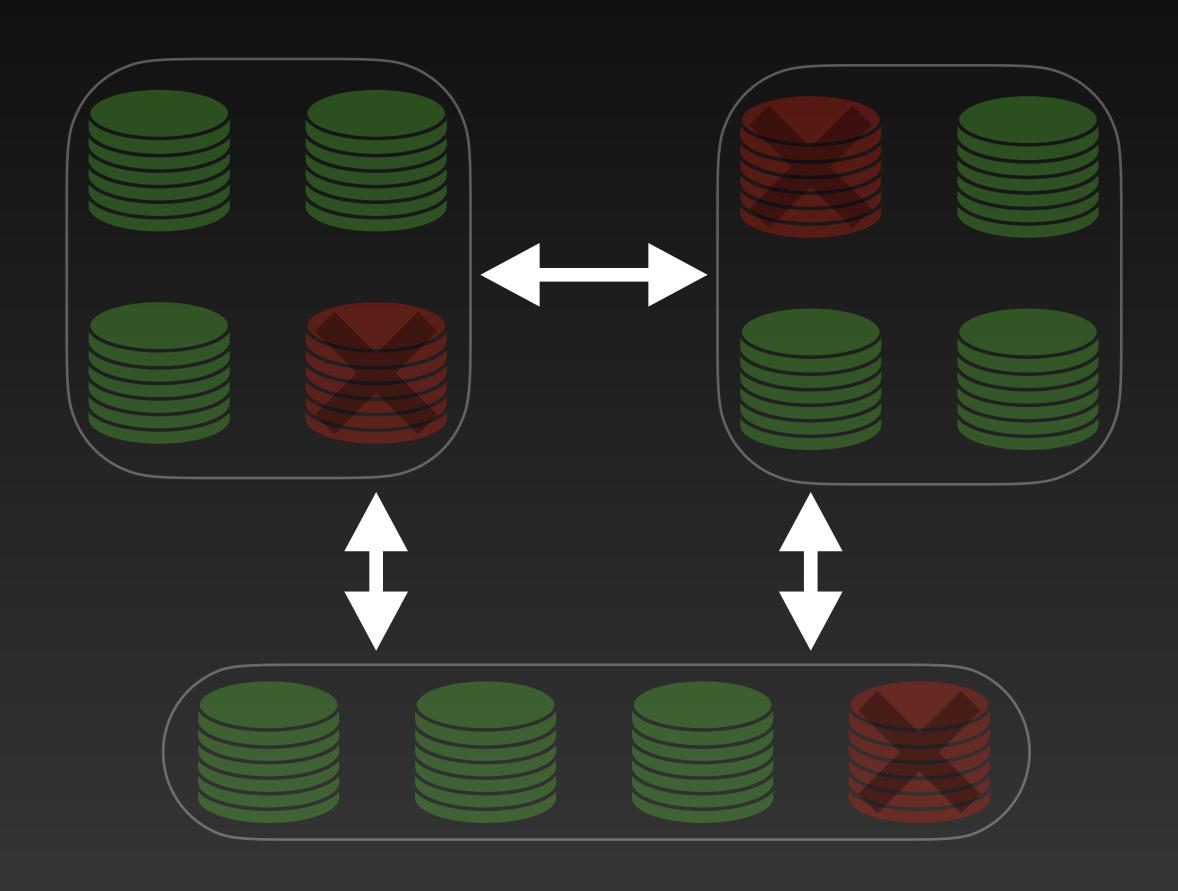


Shard 3



Cross-Shard Consensus

How do shards communicate with each other?



S-BAC

Chainspace: A Sharded Smart Contracts Platform

Mustafa Al-Bassam*, Alberto Sonnino*, Shehar Bano*, Dave Hrycyszyn † and George Danezis* * University College London, United Kingdom

Abstract—Chainspace is a decentralized infrastructure, known as a distributed ledger, that supports user defined smart contracts and executes user-supplied transactions on their objects. The correct execution of smart contract transactions is verifiable by all. The system is scalable, by sharding state and the execution of transactions, and using \mathcal{S} -BAC, a distributed commit protocol, to guarantee consistency. Chainspace is secure against subsets of infrastructure infrastructure in those parties. nodes trying to compromise its integrity or availability properties through Byzantine Fault Tolerance (BFT), and extremely highauditability, non-repudiation and 'blockchain' techniques. Even when BFT fails, auditing mechanisms are in place to trace maliious participants. We present the design, rationale, and details of Chainspace; we argue through evaluating an imple of Chamspace, we argue unough evanuating an imprimentation of the system about its scaling and other features; we illustrate a number of privacy-friendly smart contracts for smart metering, polling and banking and measure their performance.

I. INTRODUCTION

Chainspace is a distributed ledger platform for high-integrity and transparent processing of transactions within a decentralized ystem. Unlike application specific distributed ledgers, such as Bitcoin [26] for a currency, or certificate transparency [19] for certificate verification, Chainspace offers extensibility though smart contracts, like Ethereum [32]. However, users expose to infrastructure nodes: our modest testbed of 60 cores achieves 350 transactions per second, as compared with a peak rate of less than 7 transactions per second for Bitcoin over 6K full nodes. Etherium currently processes 4 transactions per second, out of theoretical maximum of 25. Furthermore, our platform is agnostic as to the smart contract language, or identity infrastructure, and supports privacy features through modern zero-knowledge techniques [3, 9].

Unlike other scalable but 'permissioned' smart contract platforms, such as Hyperledger Fabric [5] or BigchainDB [23], Chainspace aims to be an 'open' system: it allows anyone to author a smart contract, anyone to provide infrastructure on which smart contract code and state runs, and any user to access calls to smart contracts. Further, it provides ecosystem features, by allowing composition of smart contracts from different authors. We integrate a value system, named CSCoin, as a system smart contract to allow for accounting between

Network and Distributed Systems Security (NDSS) Symposium 2018 18-21 February 2018, San Diego, CA, USA ISBN 1-891562-49-5 http://dx.doi.org/10.14722/ndss.2018.23241

from traditional unpermissioned blockchains, that rely on proofof-work and global replication of state, such as Ethereum, In Chainspace smart contract authors designate the parts of the infrastructure that are trusted to maintain the integrity of their contract—and only depend on their correctness, as well as the correctness of contract sub-calls. This provides fine grained control of which part of the infrastructure need to be trusted on a per-contract basis, and also allows for horizontal scalability.

This paper makes the following contributions:

- It presents Chainspace, a system that can scale arbitrarily as the number of nodes increase, tolerates byzantine failures, and can be fully and publicly audited.
- It presents a novel distributed atomic commit protocol, called S-BAC, for sharding generic smart contract transactions across multiple byzantine nodes, and correctly coordinating those nodes to ensure safety, liveness and security properties.
- It introduces a distinction between parts of the smart contract that execute a computation, and those that check the computation and discusses how that distinction is key to supporting privacy-friendly smart-
- It provides a full implementation and evaluates the per-formance of the byzantine distributed commit protocol, S-BAC, on a real distributed set of nodes and under
- It presents a number of key system and application smart contracts and evaluates their performance The contracts for privacy-friendly smart-metering and privacy-friendly polls illustrate and validate support for high-integrity and high-privacy applications.

Outline: Section II presents an overview of Chainspace Section III presents the client-facing application interface; Section IV presents the design of internal data structures guaranteeing integrity, the distributed architecture, the byzantine commit protocols, and smart contract definition and composition. Section V argues the correctness and security; specific smart contracts and their evaluations are presented in Section VI Section VII presents an evaluation of the core protocols and smart contract performance; Section VIII presents limitation and Section IX a comparison with related work; and Section X

II. SYSTEM OVERVIEW

Chainspace allows applications developers to implement distributed ledger applications by defining and calling proce-

NDSS'18

Atomix

OmniLedger: A Secure, Scale-Out, Decentralized Ledger via Sharding

Abstract-Designing a secure permissionless distributed ledger (blockchain) that performs on par with centralized payment processors, such as Visa, is a challenging task. Most existing distributed ledgers are unable to scale-out, i.e., to grow their total processing capacity with the number of validators; and those that o, compromise security or decentralization. We present Om niLedger, a novel scale-out distributed ledger that preserves long term security under permissionless operation. It ensures security and correctness by using a bias-resistant public-randomness protocol for choosing large, statistically representative shards that process transactions, and by introducing an efficient crossshard commit protocol that atomically handles transactions affecting multiple shards. OmniLedger also optimizes performance via parallel intra-shard transaction processing, ledger pruning collectively-signed state blocks, and low-latency "trust-but verify" validation for low-value transactions. An evaluation of our experimental prototype shows that Onnii-ledger's throughput scales linearly in the number of active validators, supporting

transaction volume and the number of independent participants involved in processing them, is a major challenge to their mainstream adoption, especially when weighted against security and decentralization challenges. Many approaches transaction volume and the number of independent particiexhibit different security and performance trade-offs [10], [21], [32], [40]. Replacing the Nakamoto consensus [36] as its validator set. To support the more power-efficient alwith PBFT [13], for example, can increase throughput while decreasing transaction commit latency [1], [32]. These approaches still require all *validators* or consensus group membership based on directly invested stake rather than work, OmniLedger builds on Ouroboros [31] and Algorand [25], running a public bers to redundantly validate and process all transactions, hence the system's total transaction processing capacity does

databases, whose capacity scales horizontally with the number of participants, is by *sharding* [14], or partitioning the state t-of-n threshold assumptions. into multiple shards that are handled in parallel by different subsets of participating validators. Sharding could benefit OmniLedger addresses the second key security challenge of DLs [15] by reducing the transaction processing load on each validator and by increasing the system's total processing capacity proportionally with the number of participants. Existing

OmniLedger chooses shards large enough, based on the analproposals for sharded DLs, however, forfeit permissionless ysis in Section VI, to ensure a negligible probability that any decentralization [16], introduce new security assumptions, shard is ever compromised, even across years of operation. and/or trade performance for security [34], as illustrated in Finally, to ensure that transactions either commit or abort atomically even when they affect state distributed across multi-

provides "scale-out" transaction processing capacity competitive with centralized payment-processing systems, such as



faces three key correctness and security challenges. First, OmniLedger must choose statistically representative groups of validators periodically via permissionless Sybil-attackour experimental prototype shows that Ommitedget's unoughpost scales linearly in the number of active validators, supporting Visa-level workloads and beyond, while confirming typical trans-cretions in under two seconds.)forming shards (subsets of validators to record state and process transactions), that are both sufficiently large and bias-The scalability of distributed ledgers (DLs), in both total ransaction volume and the number of independent participation volume and the number of independent participations.

not increase with added participants, and, in fact, gradually decreases due to increased coordination overheads.

The proven and obvious approach to building "scale-out" the current stakeholder distribution defined in the ledger. To ensure that this sampling of representative validators is both scalable and strongly bias-resistant, OmniLedger uses

Figure 1 and explored in detail in Sections II and IX.

We introduce OmniLedger, the first DL architecture that

We introduce OmniLedger, the first DL architecture that Visa, without compromising security or support for permissionless decentralization. To achieve this goal, OmniLedger and unlock state affected by partially completed transactions.

S&P'18

Cross-Shard Consensus

Byzantine Agreement



2-Phases Atomic Commit

Spoiler alert: Insecure under parallel composition

S-BAC

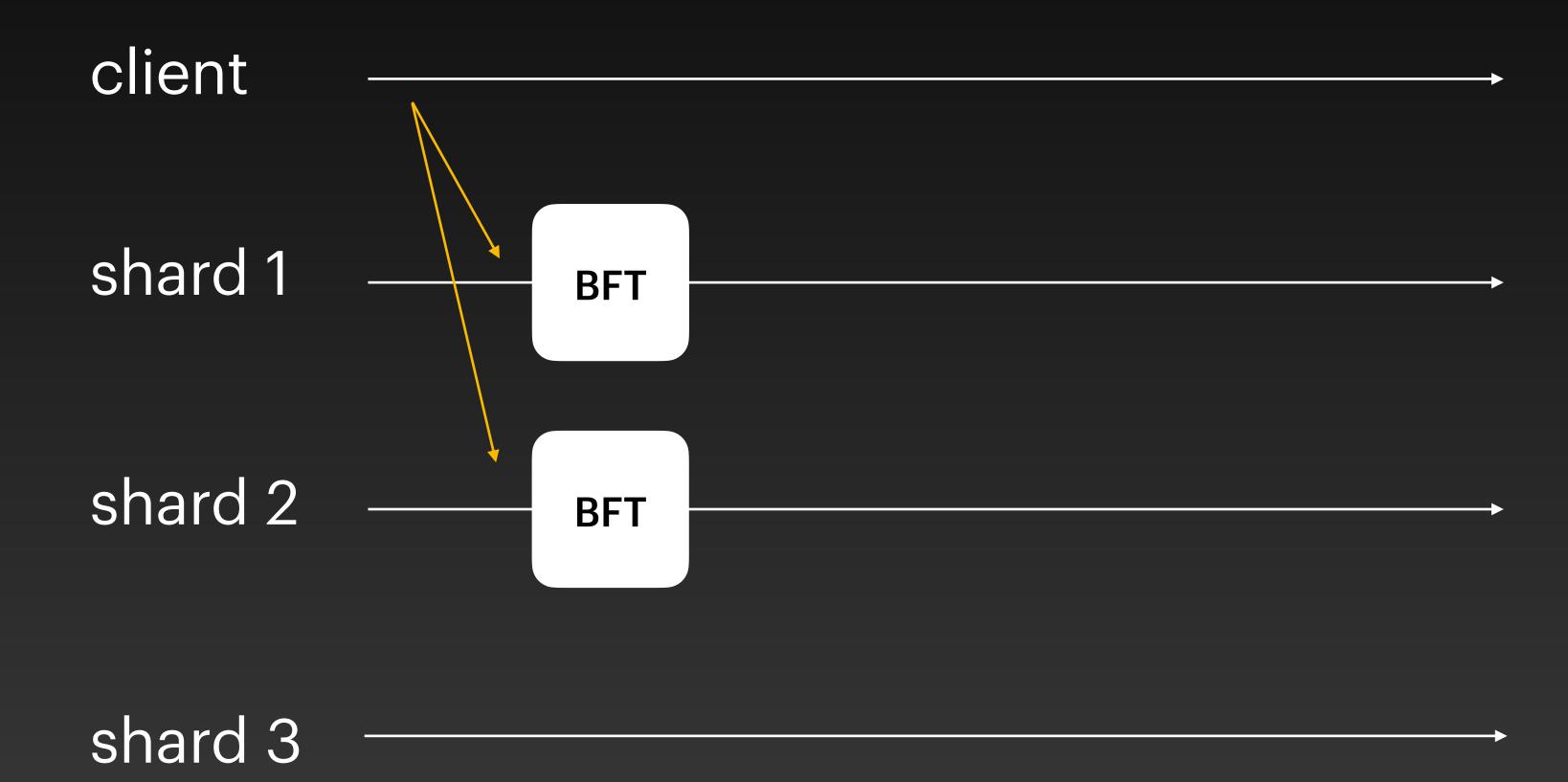
$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$

client —

shard 1 _____

shard 3

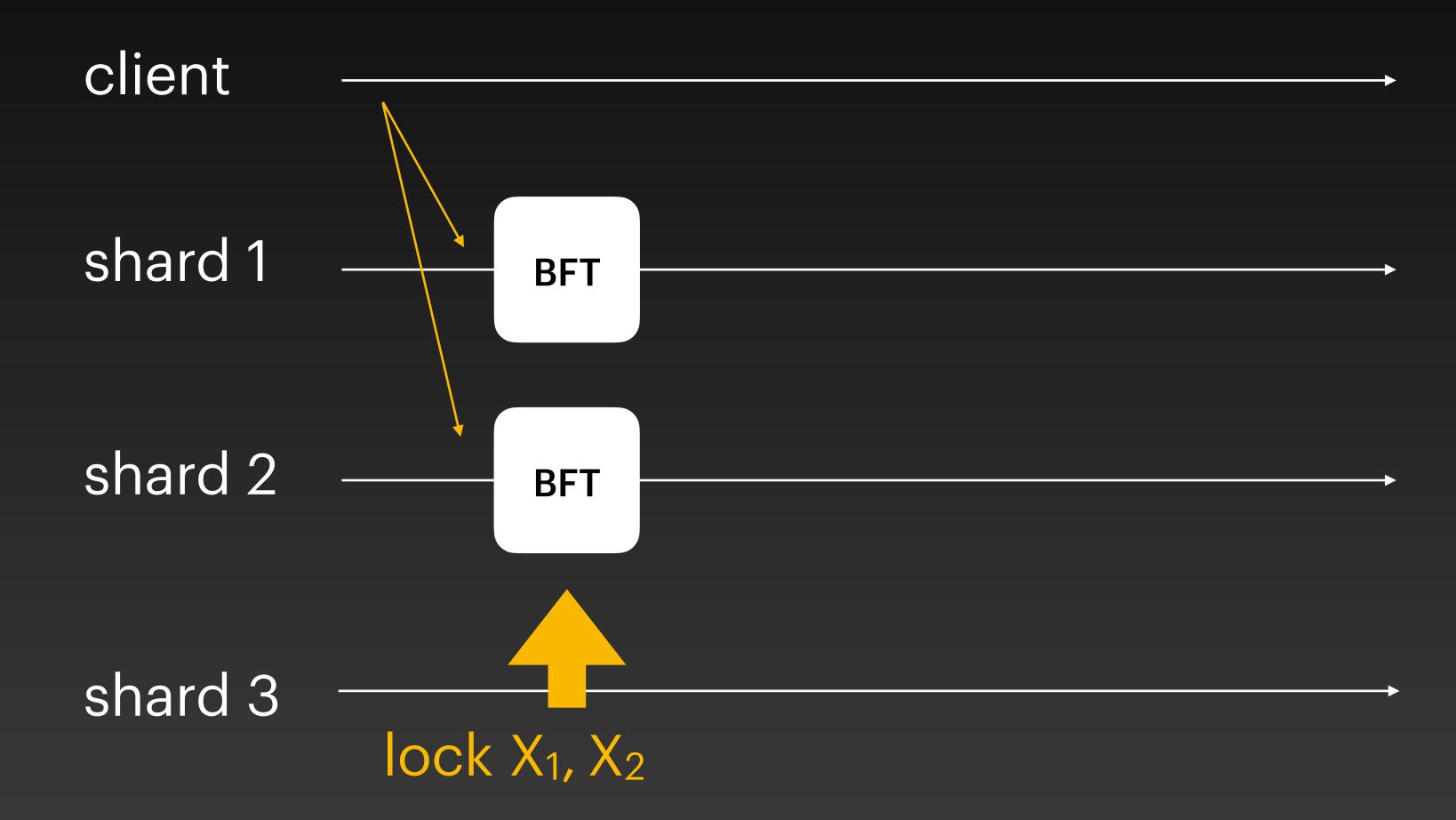
$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$



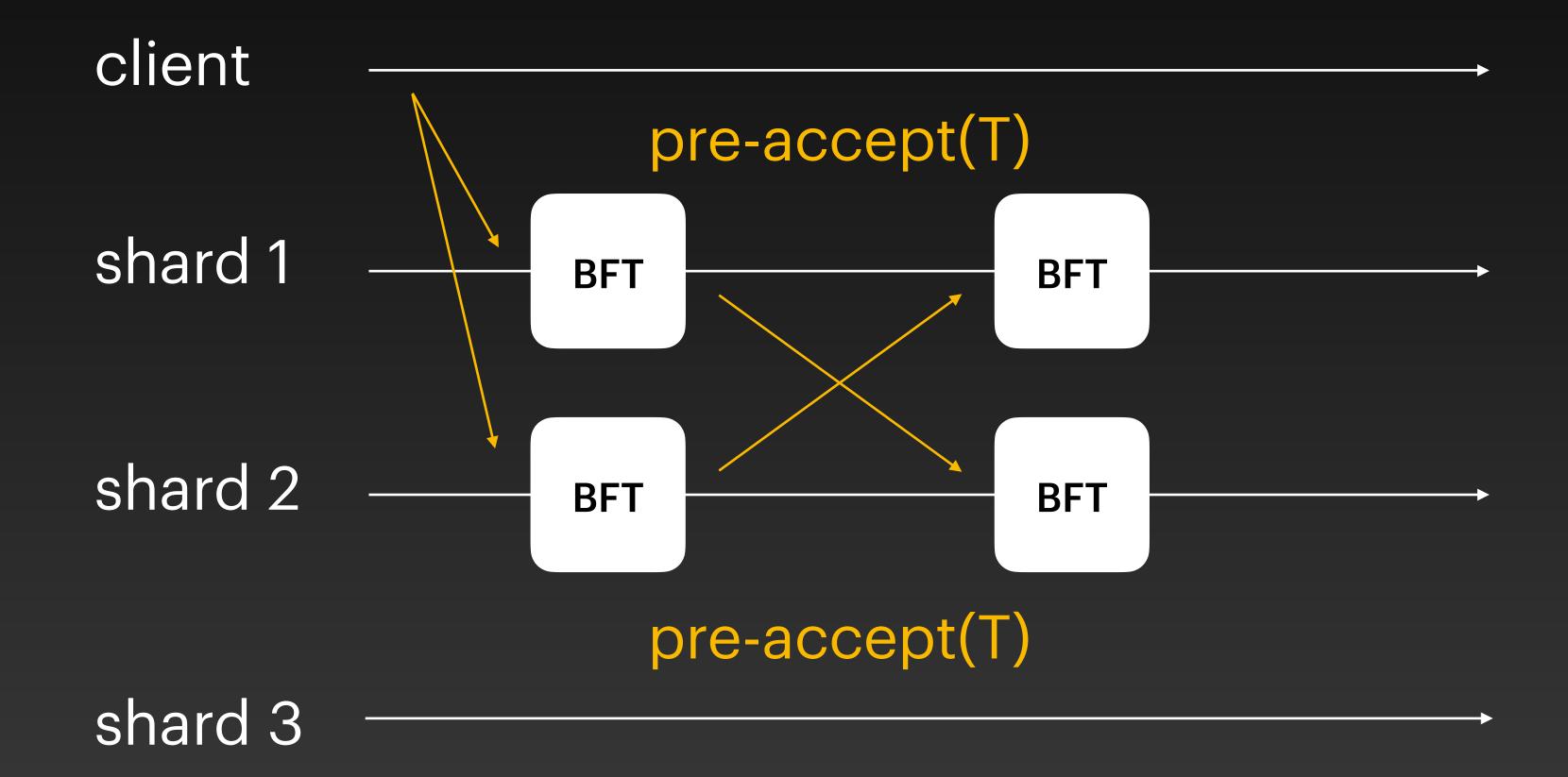
S-BAC
$$T(x_1, x_2) \to (y_1, y_2, y_3)$$



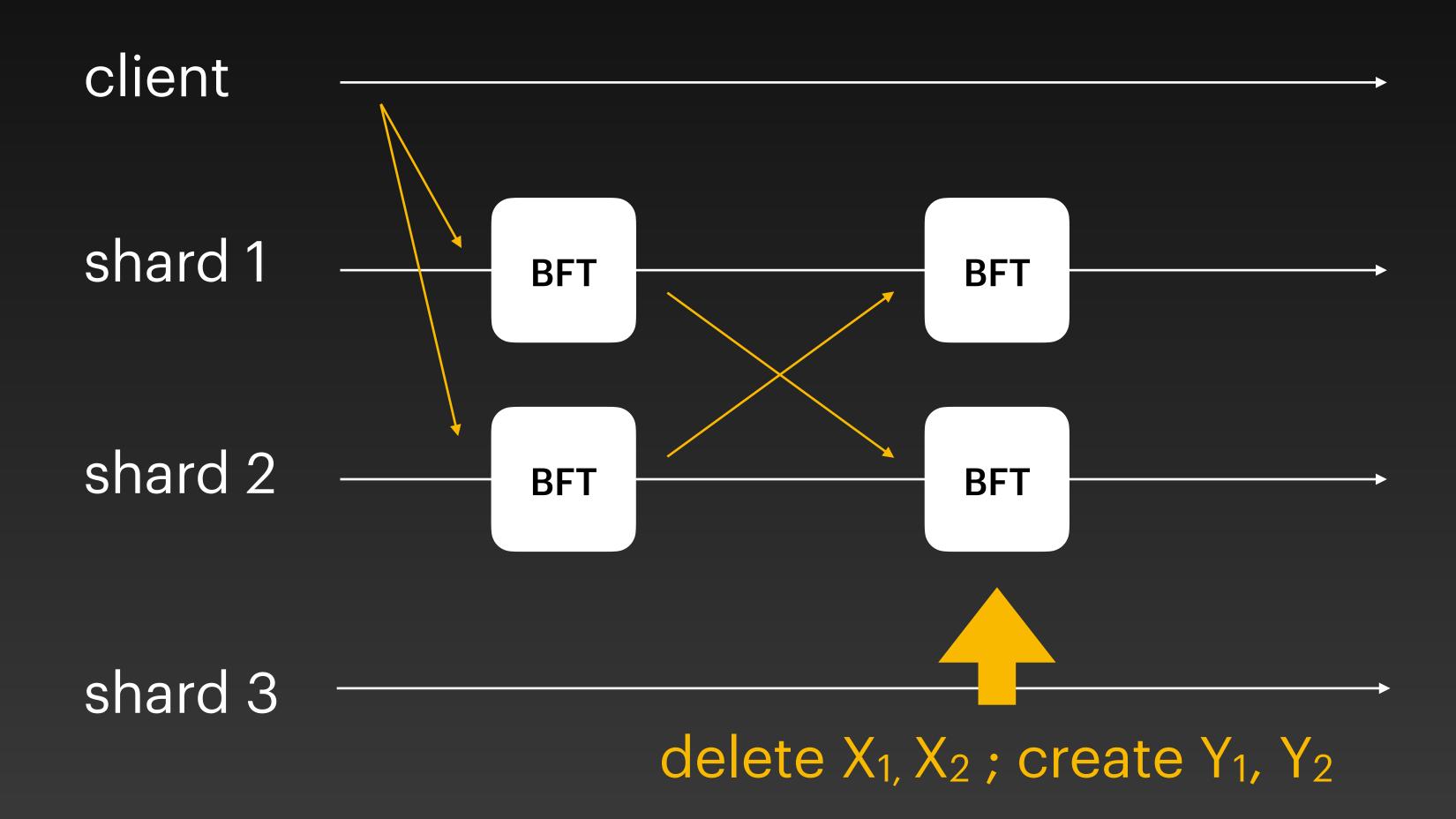
$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$



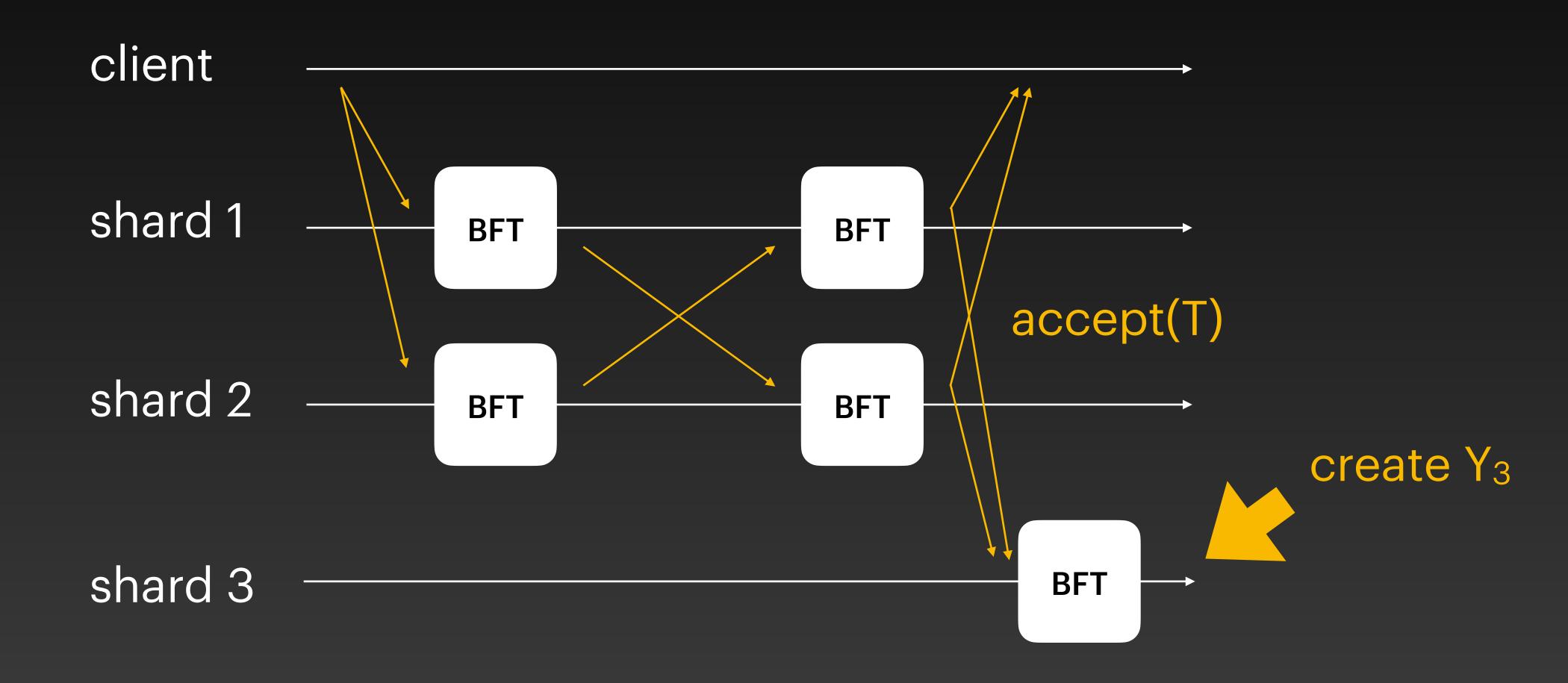
$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$



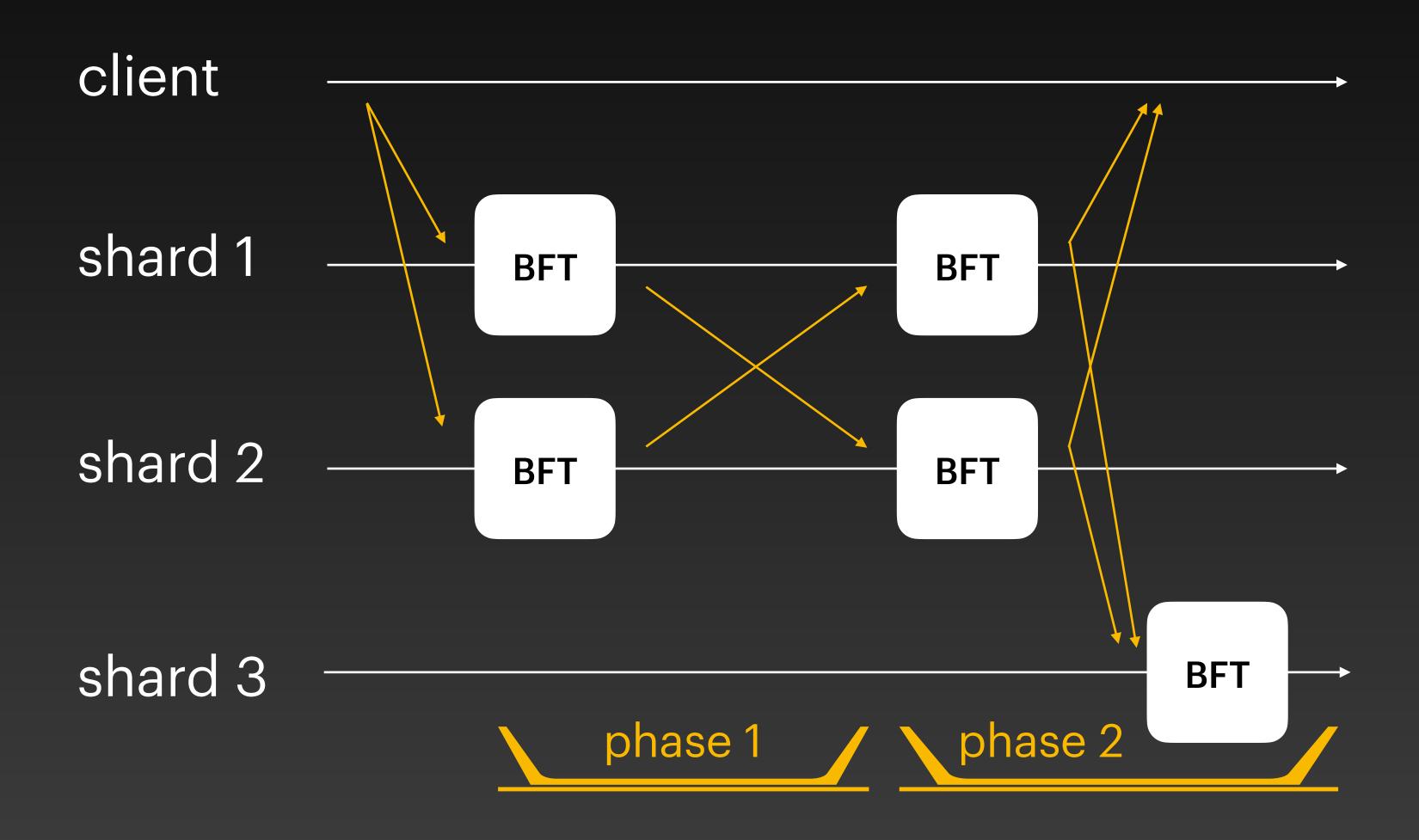
$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$



$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$

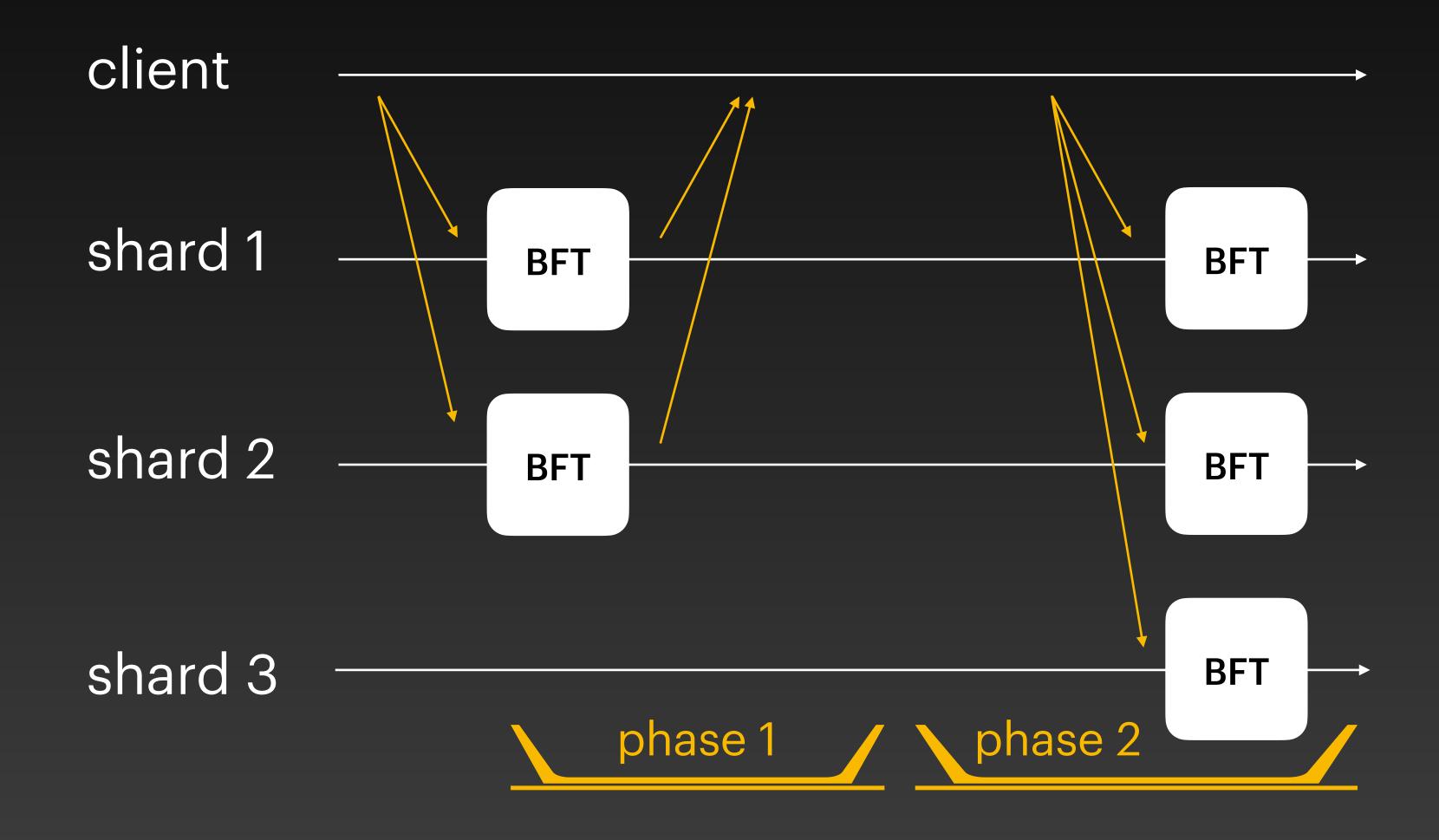


$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$



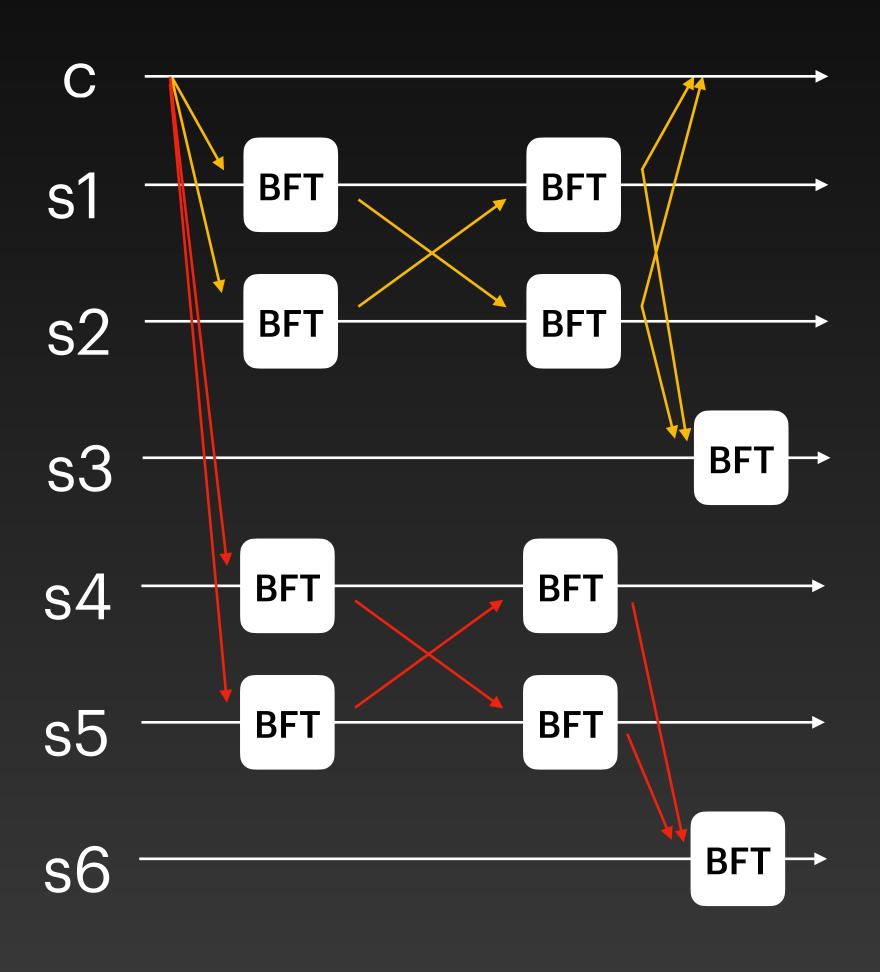
Atomix

$$T(x_1, x_2) \rightarrow (y_1, y_2, y_3)$$



Cross-Shard Consensus

How does it achieve linear scalability?

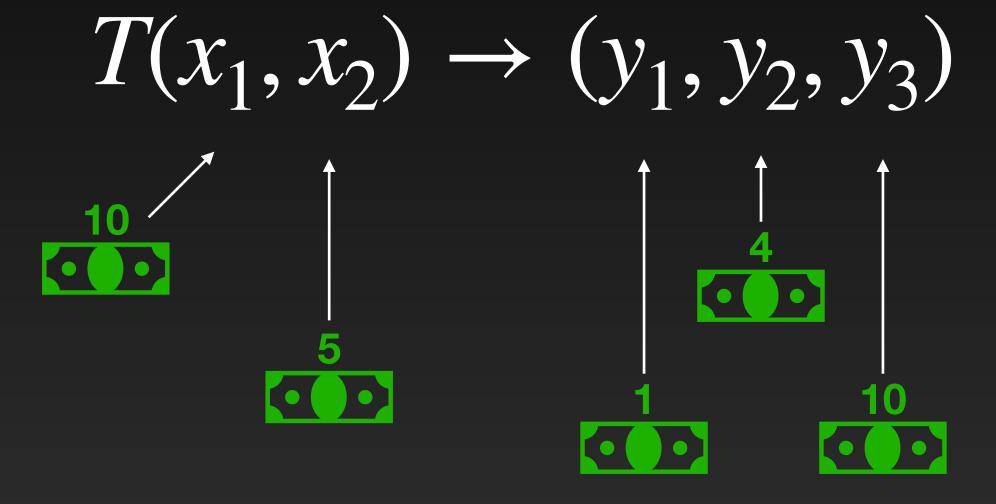


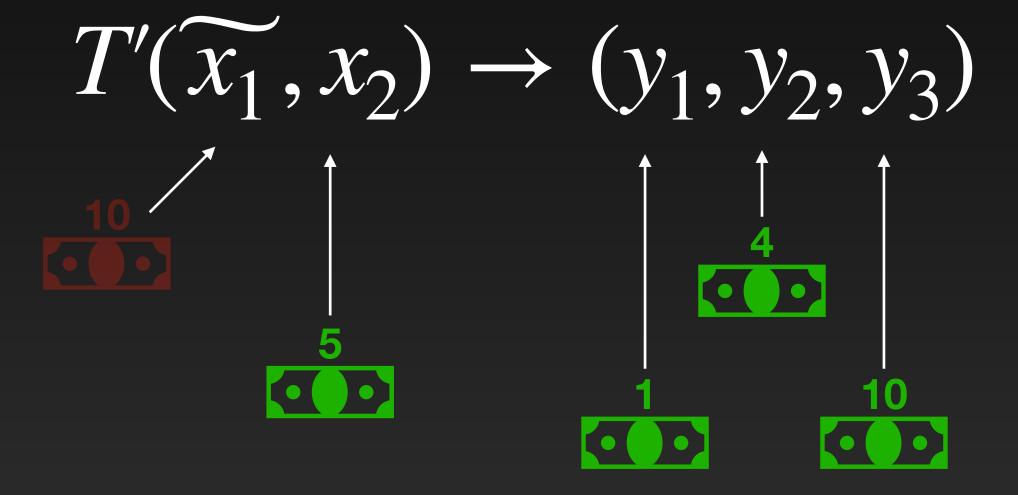
Insecure under parallel composition

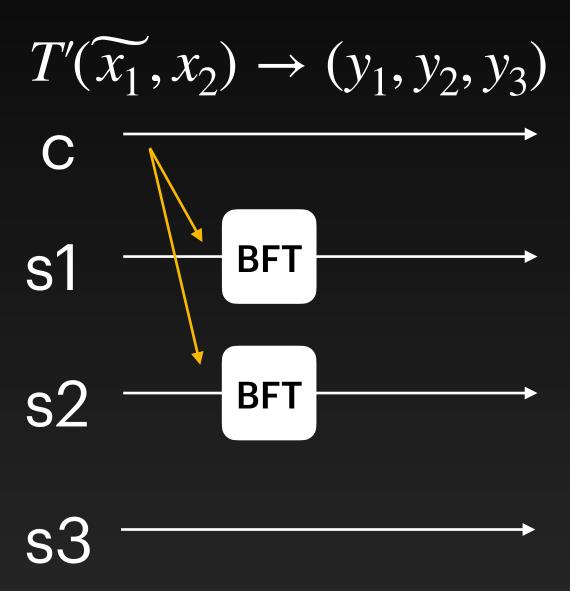
Attacks

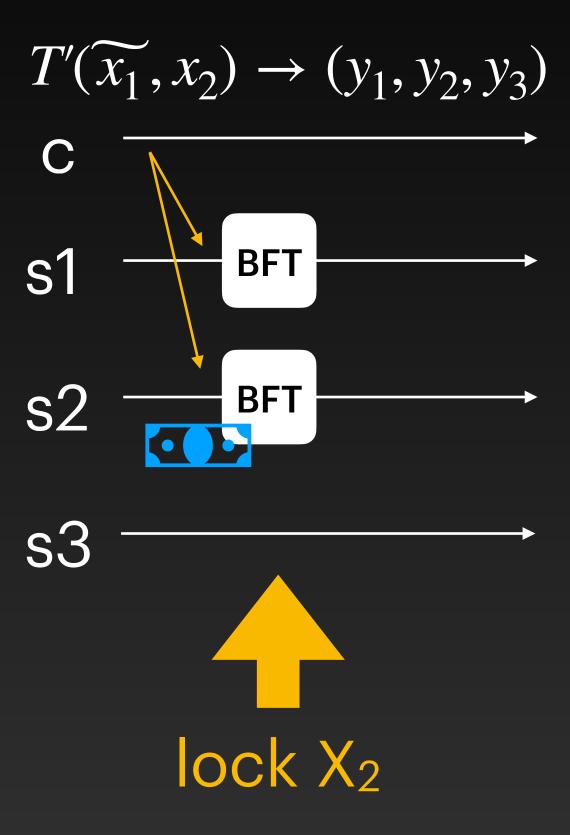
Double spend any object

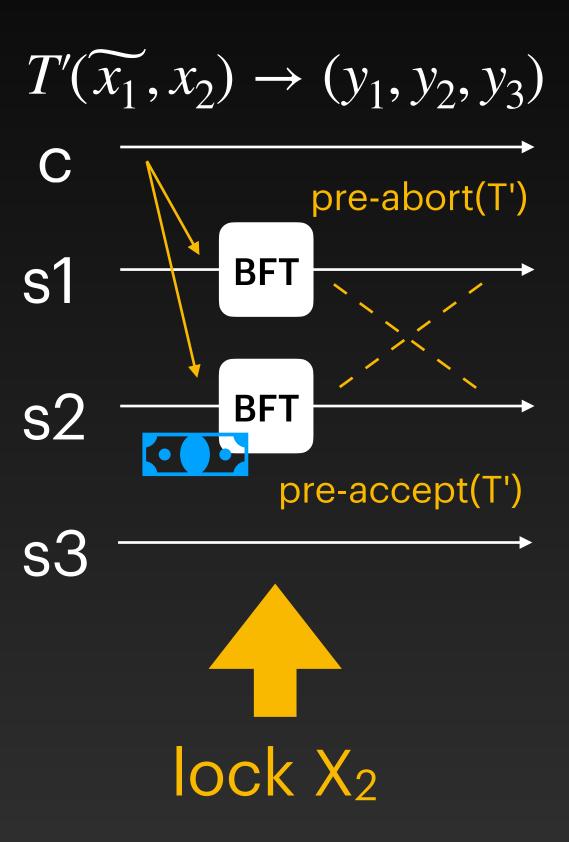
- Does not need to collude with any node
- Acts as client or passive observer
- Re-orders network messages (not always needed)

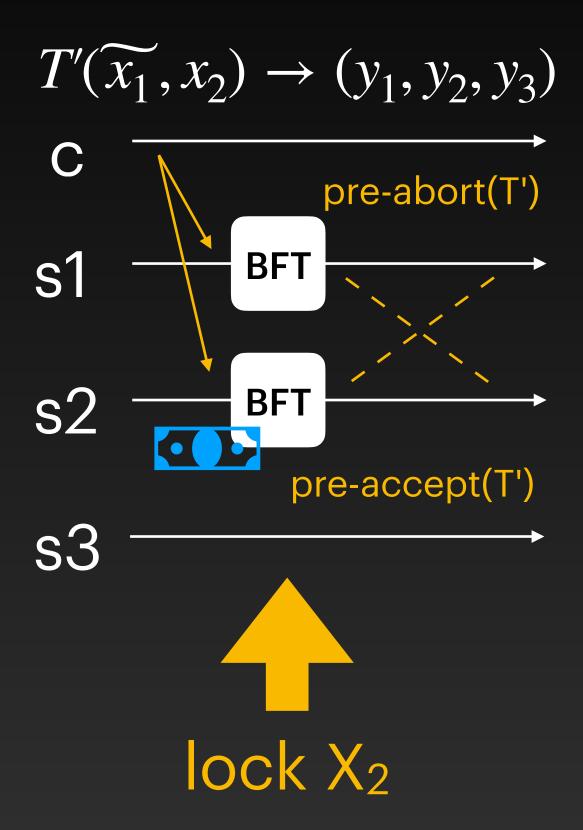


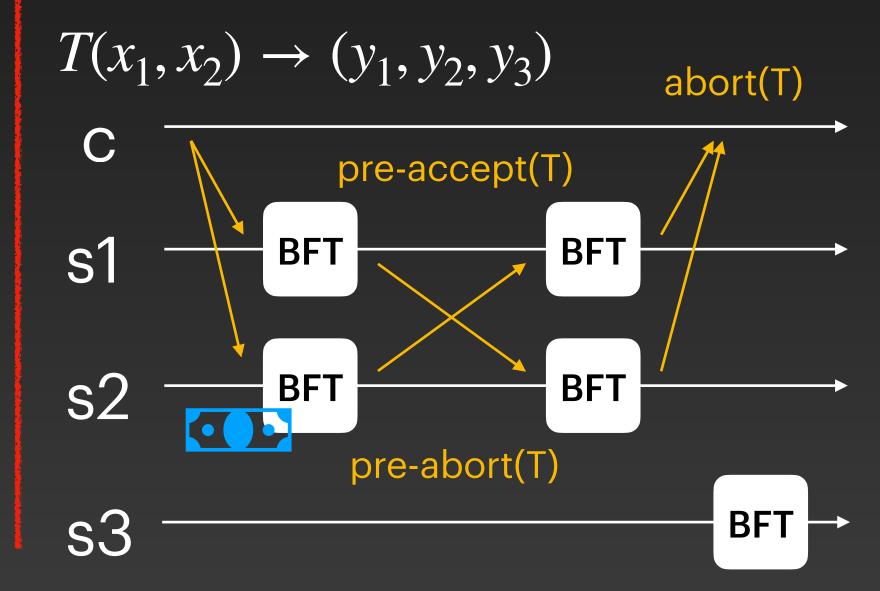


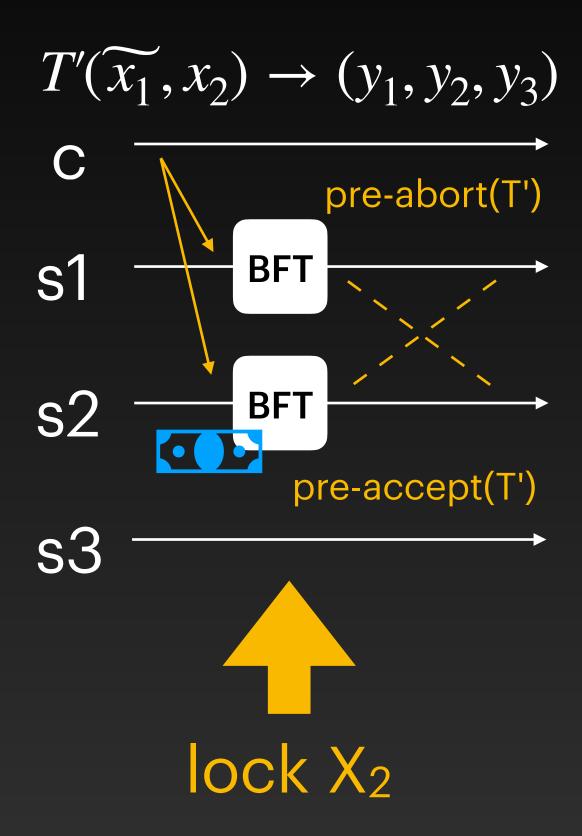




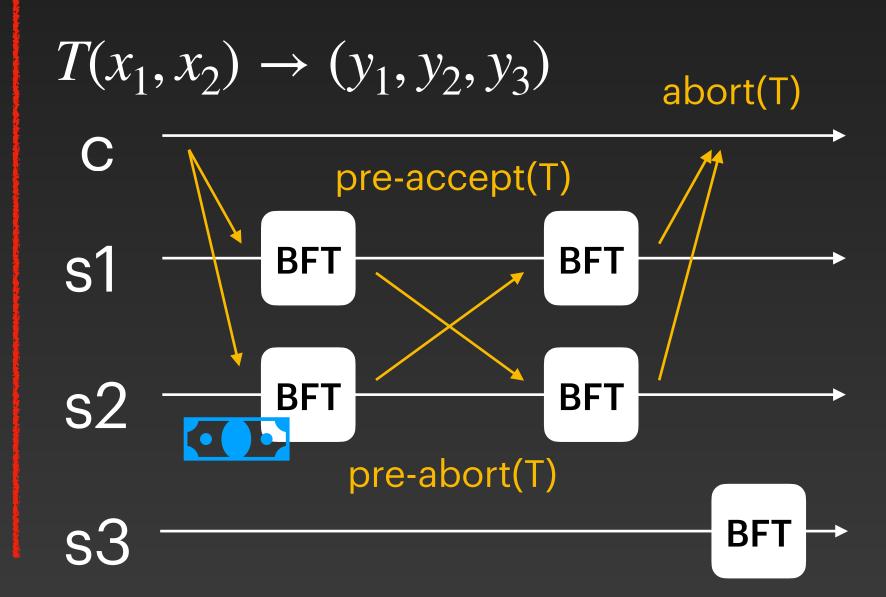


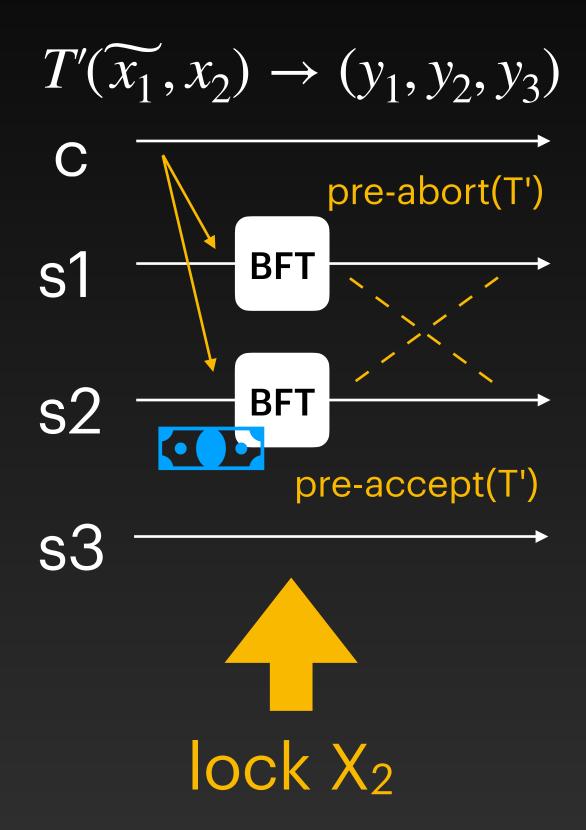


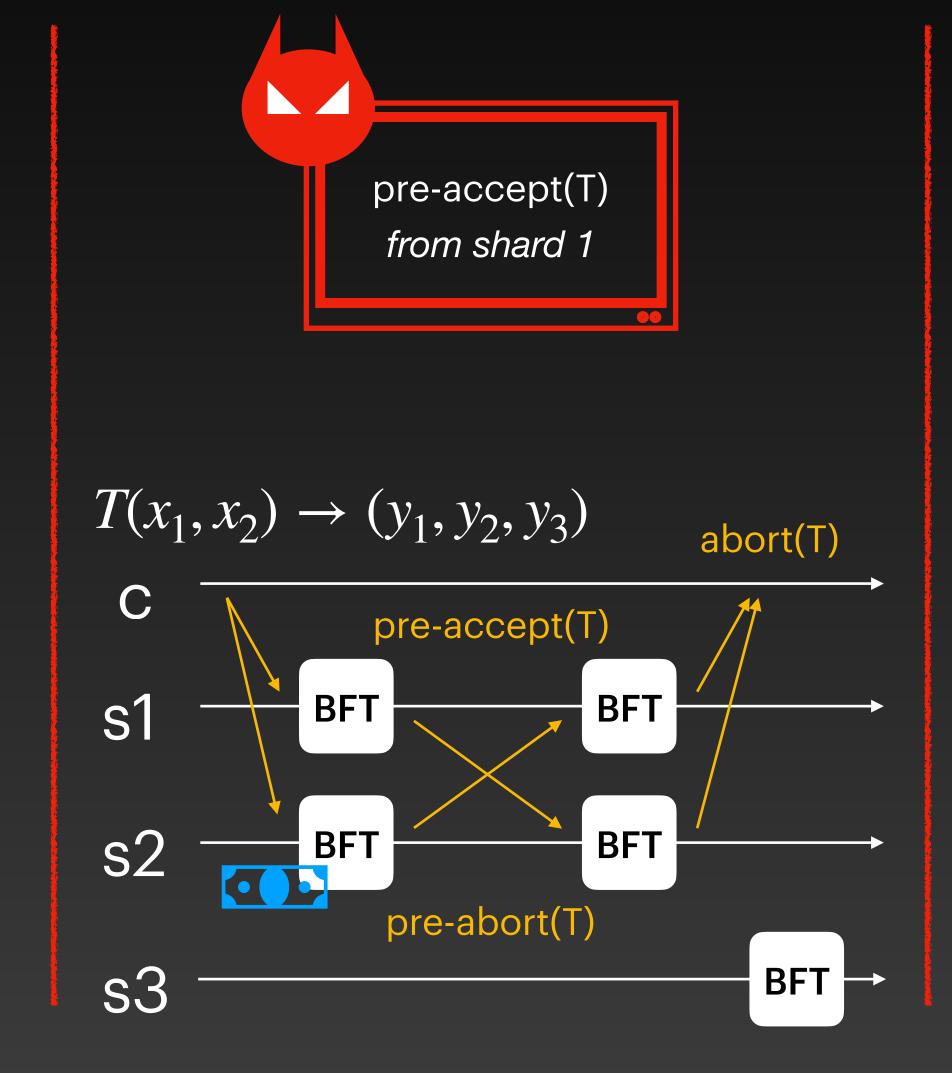


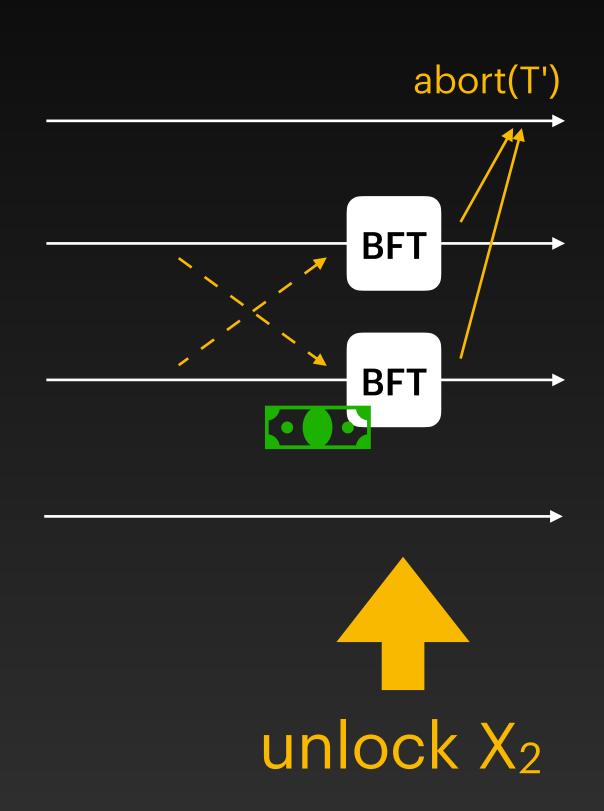






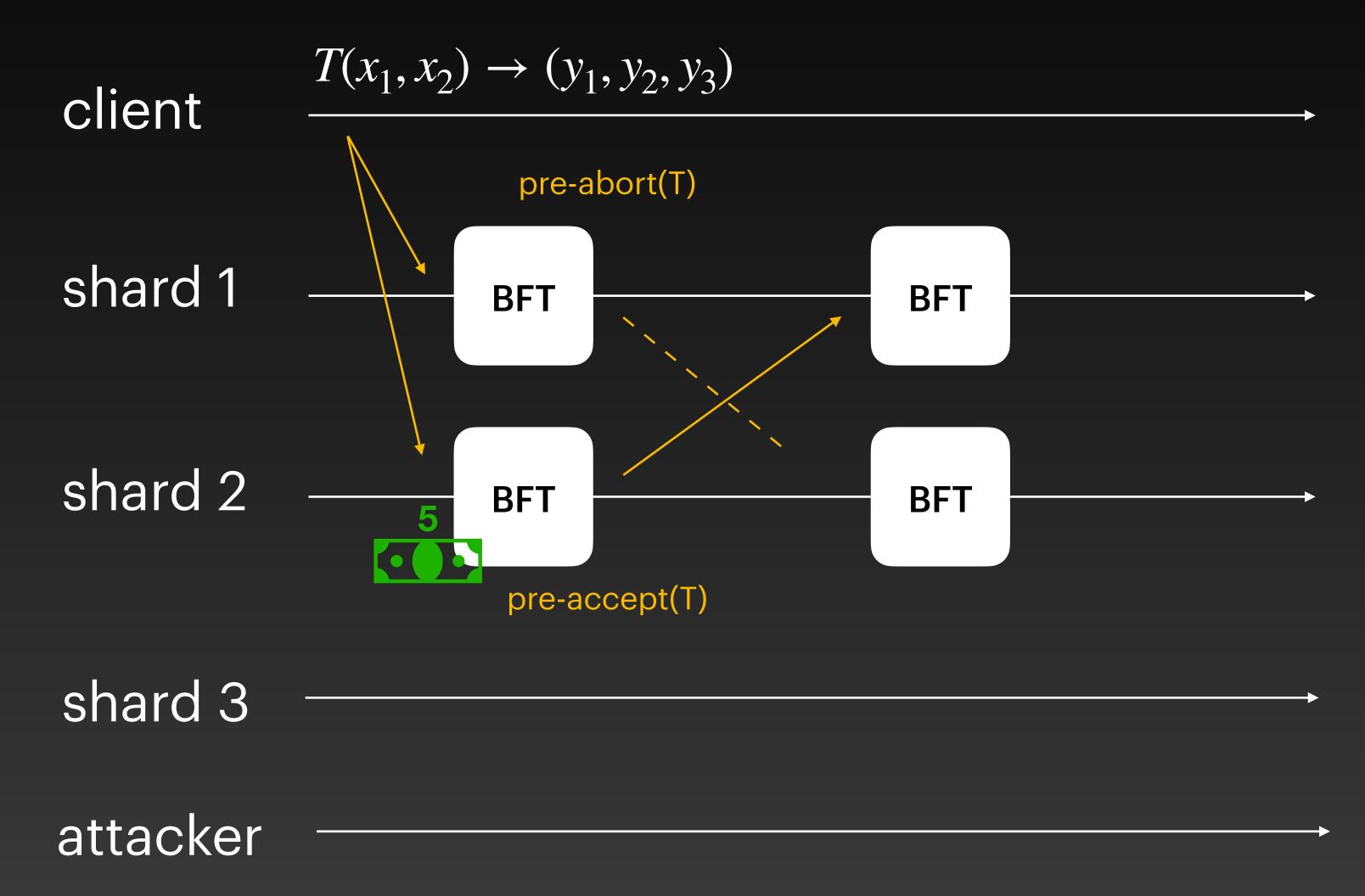


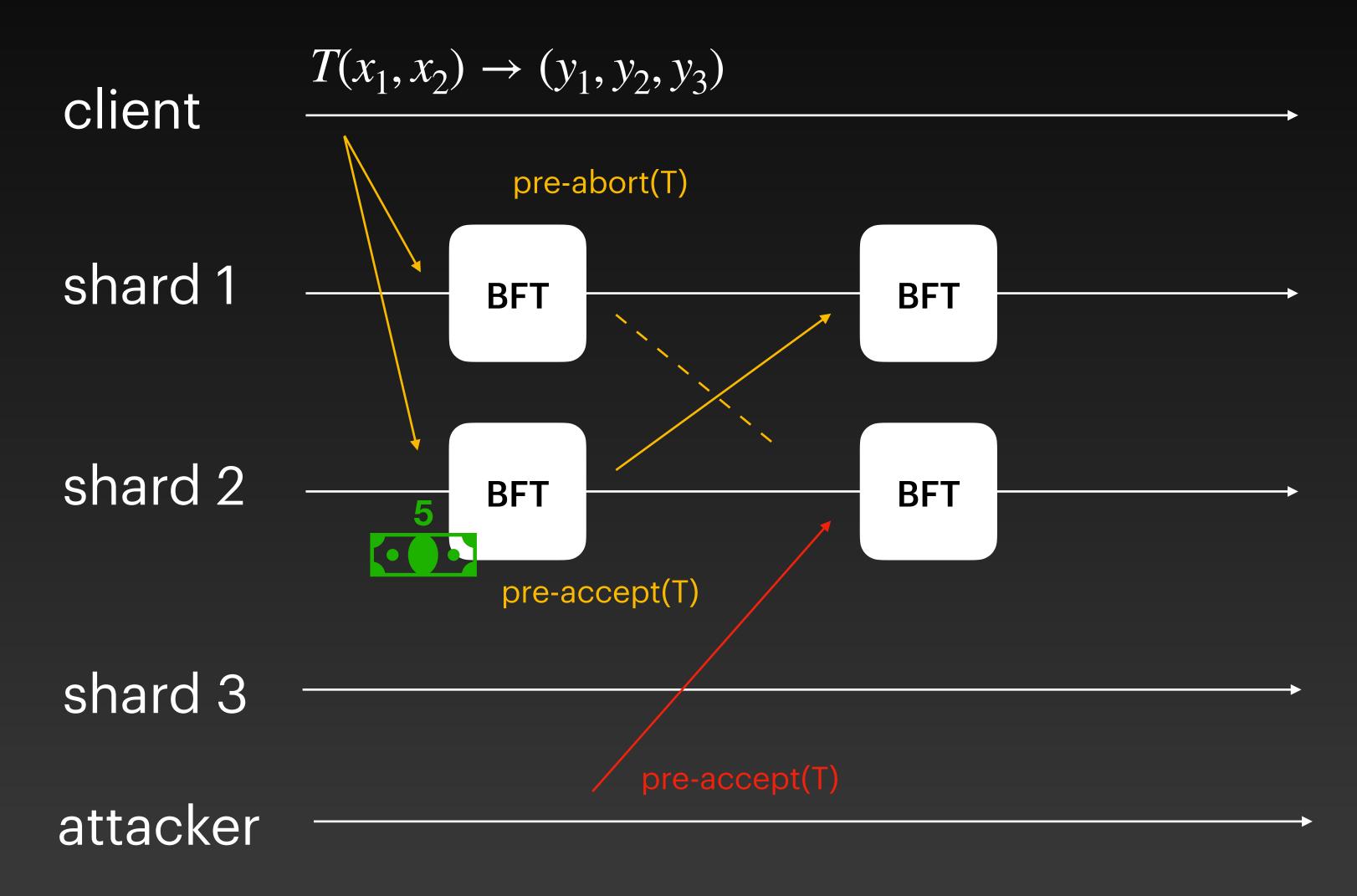




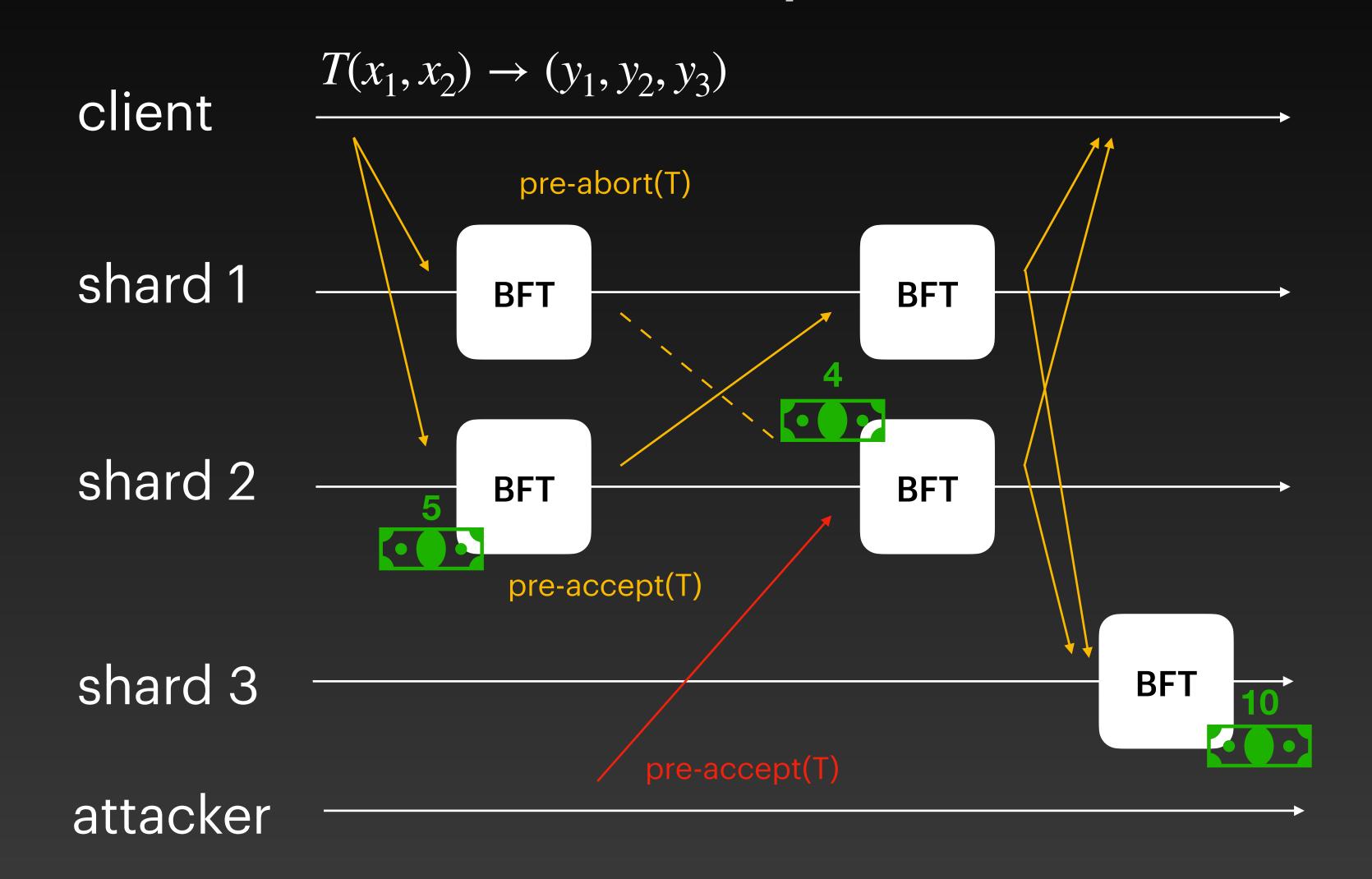
Attack against S-BAC Double-spend X₁

$$T^*(x_1) \to (y_*)$$
client
$$Shard 1$$
BFT
$$10$$



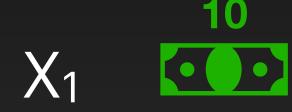


Attack against S-BAC Double-spend X₁



Double-spend X₁

Before attack



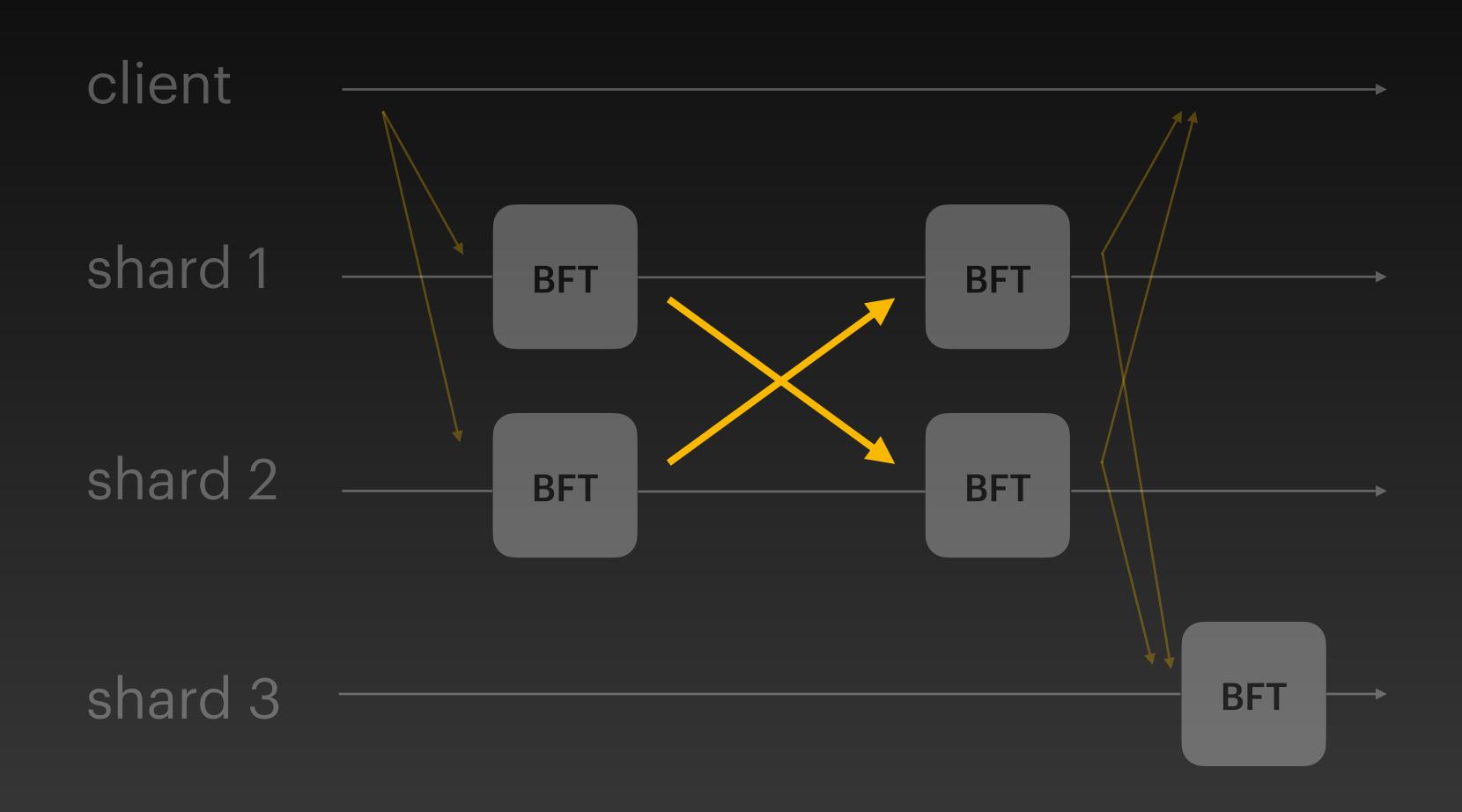
$$\chi_2$$

After attack



If it is not implemented, it does not work

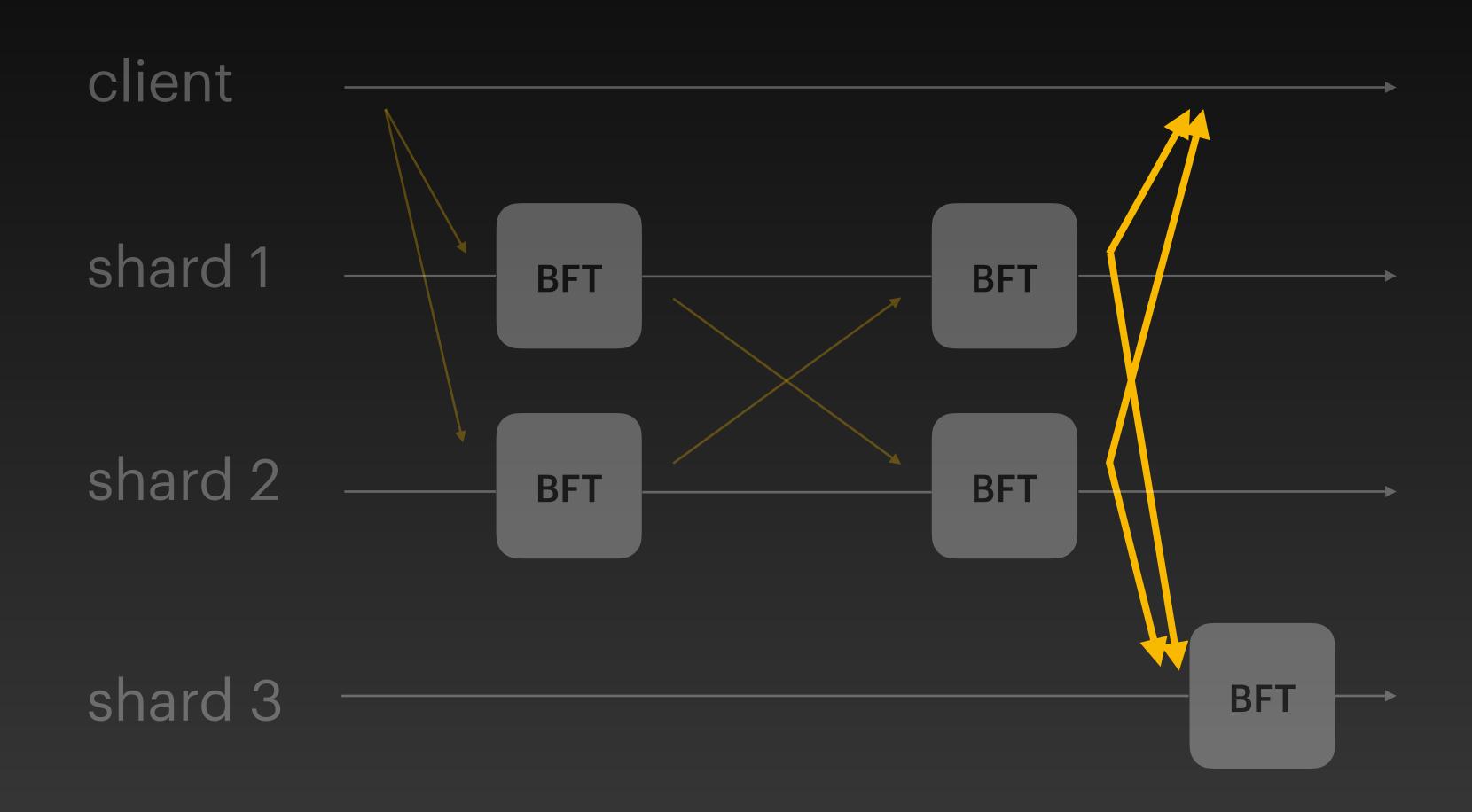
Attacks against S-BAC First phase



Attacks against S-BAC First phase

	Phase 1 of S-BAC		Phase 2 of S-BAC		
	Shard 1 (potential victim)	Shard 2 (potential victim)	Shard 1 (potential victim)	Shard 2 (potential victim)	Shard 3 (potential victim)
1	$ pre-accept(T) \\ lock x_1 $	$ pre-accept(T) \\ lock x_2 $	accept(T) create y_1 ; inactivate x_1	accept(T) create y_2 ; inactivate x_2	- create y ₃
2	⊳pre-abort(T)		accept(T) create y_1 ; inactivate x_1	$abort(T)$ $unlock x_2$	- create y ₃
3		⊳pre-abort(T)	$abort(T)$ $unlock x_1$	accept(T) create y_2 ; inactivate x_2	- create y ₃
4	⊳pre-abort(<i>T</i>)	⊳pre-abort(<i>T</i>)	$abort(T)$ unlock x_1	$abort(T)$ unlock x_2	-
5	pre-abort(T) -	pre-accept(T) lock x_2	abort(T)	abort(T) unlock x_2	-
6	⊳pre-accept(T)		abort(T)	accept(T) create y_2 ; inactivate x_2	- create y ₃
7	pre-accept(T) lock x_1	pre-abort(T) -	$\begin{array}{c} abort(T) \\ unlock\ x_1 \end{array}$	abort(T) -	-
8		$\triangleright pre\text{-}accept(T)$	accept(T) create y_1 ; inactivate x_1	abort(T)	- create y ₃
9	pre-abort(T)	pre-abort(T)	abort(T)	abort(T)	-

Attacks against S-BAC Second phase



Attacks against S-BAC Second phase

	Phase 2 of S-BAC					
	Shard 1	Shard 2	Shard 3 (potential victim)			
1	accept(T)	accept(T)	_			
	create y_1 ; inactivate x_1	create y_2 ; inactivate x_2	create y ₃			
2	\triangleright accept (T)		create y ₃			
3		$\triangleright accept(T)$	create y ₃			
1	$\triangleright accept(T)$	$\triangleright accept(T)$	create y ₃			
5	abort(T)	abort(T)	_			
	$(unlock x_1)$	$(\text{unlock } x_2)$	_			
5	\triangleright accept (T)		create y ₃			
7		$\triangleright accept(T)$	create y ₃			
3	\triangleright accept (T)	$\triangleright accept(T)$	create y ₃			

What causes these issues?

Issue 1. Input shards cannot associate protocol messages to a specific protocol execution.

Issue 2. Output shards (that are not also input shards) do not experience the first phase of the protocol

Easy Fix?

Global sequence numbers?

Wait for messages to arrive?

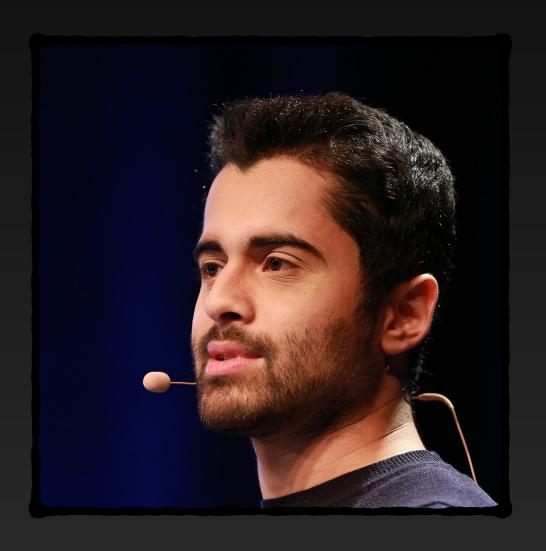




Byzcuit

S-BAC + Atomix

Byzcuit Acknowledgments



Mustafa Al-Bassam



Alberto Sonnino



Bano Shehar

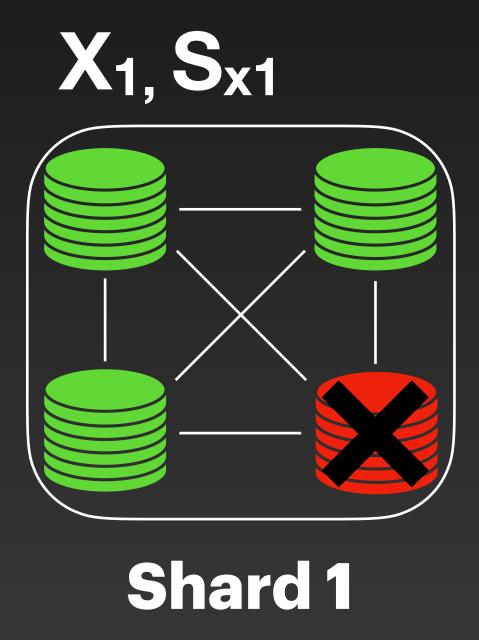


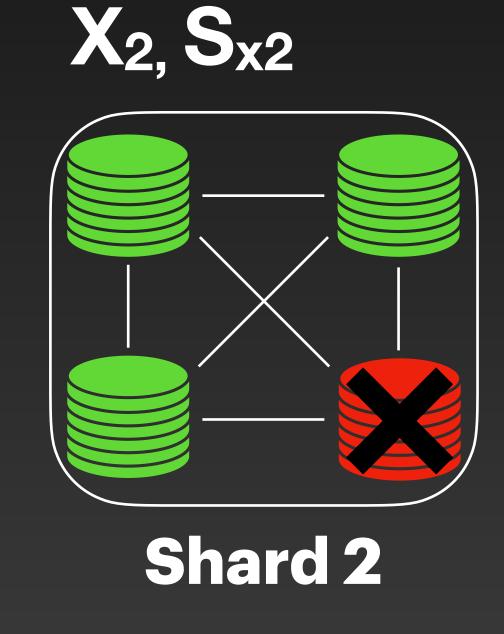
George Danezis

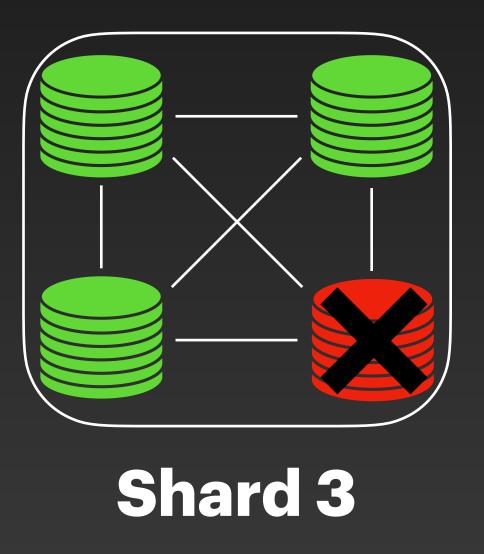
University College London

Byzcuit Fix issue 1

Add sequence numbers per object

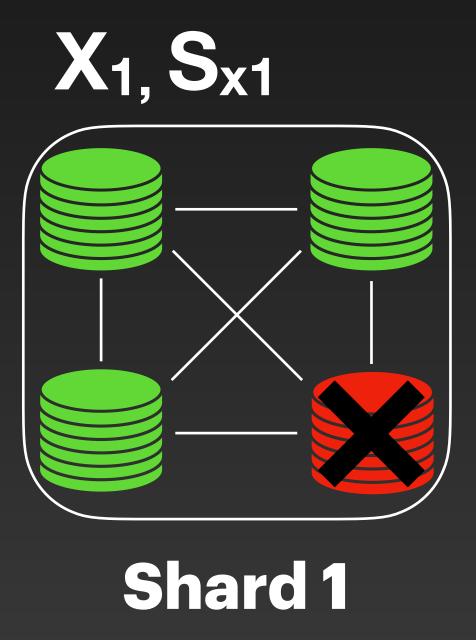


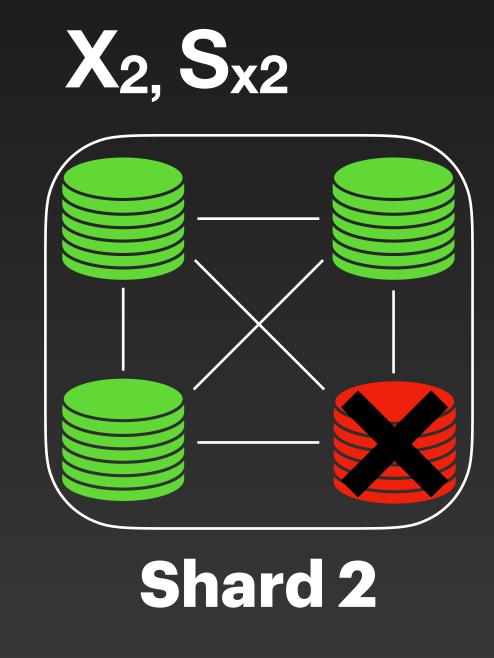


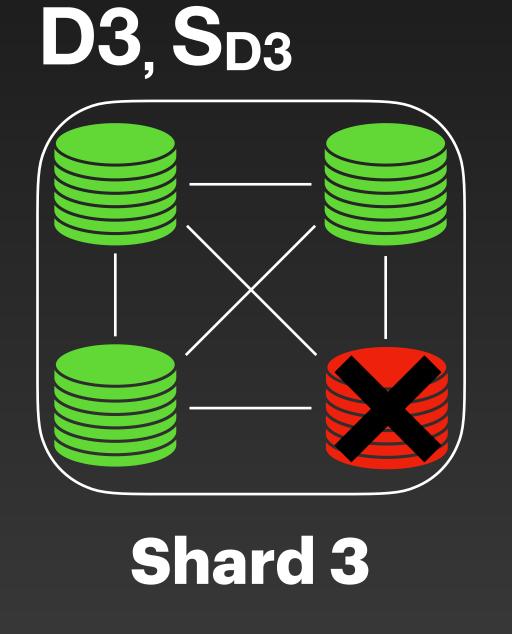


Byzcuit Fix issue 2

Dummy objects for output shards

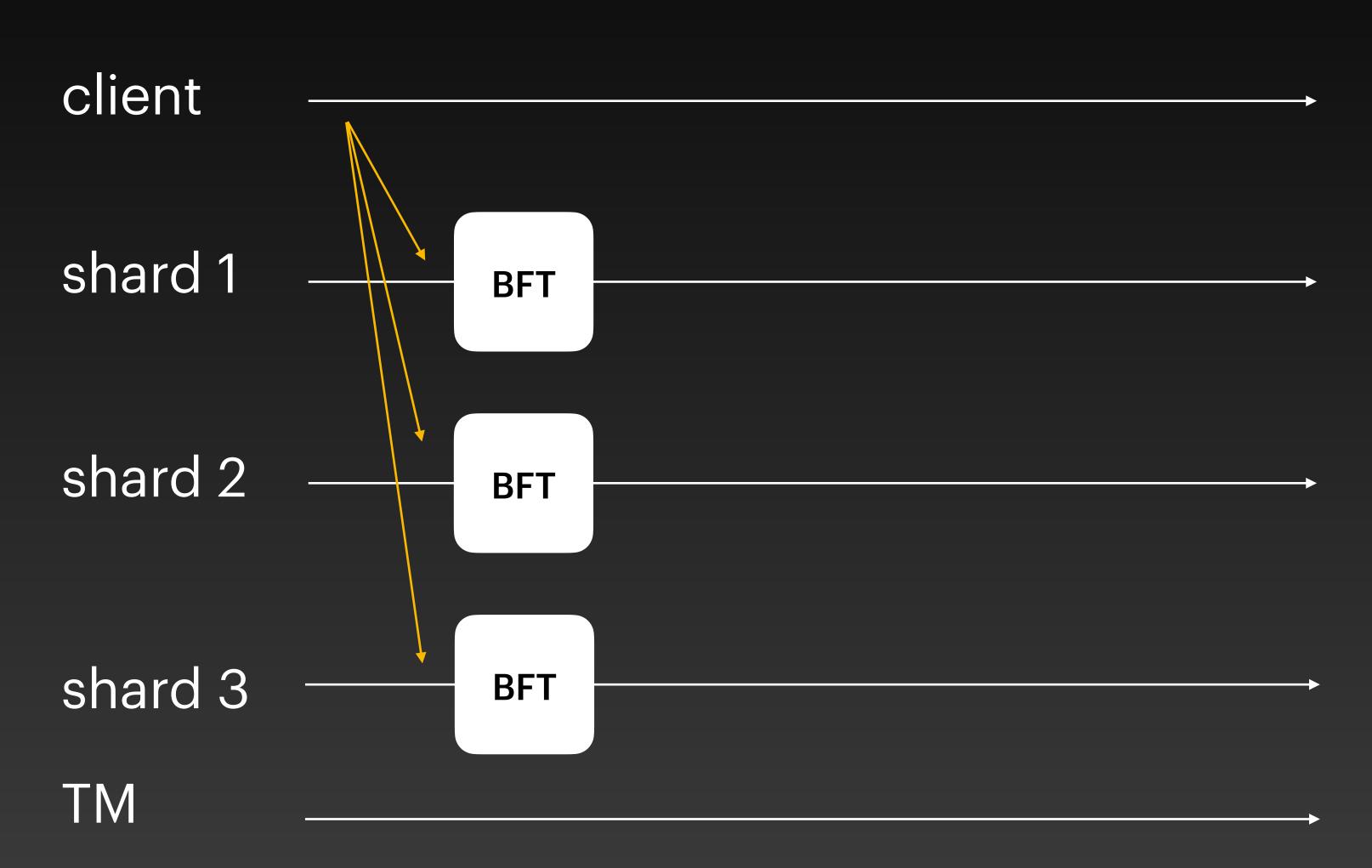




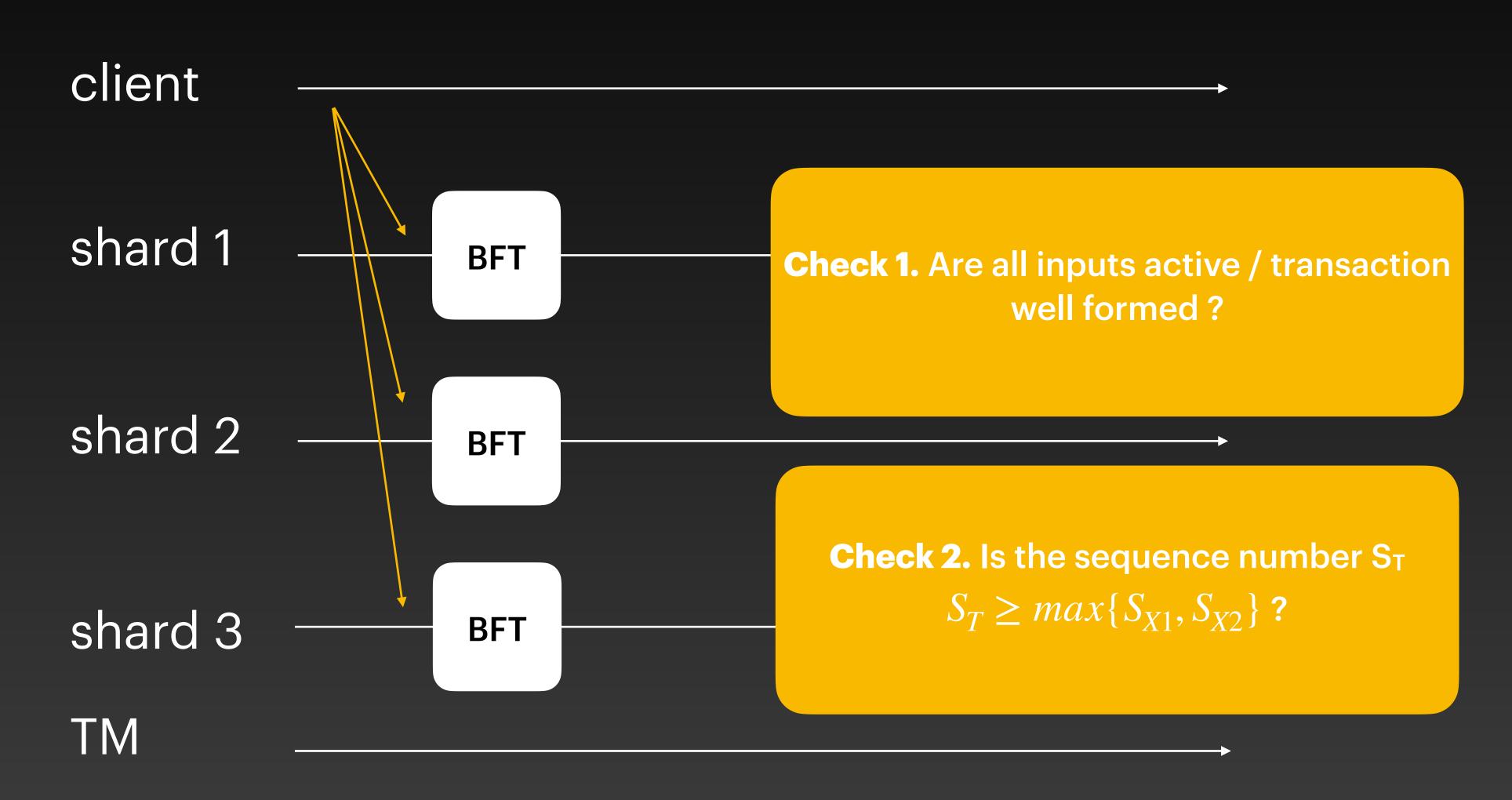


Byzcuit

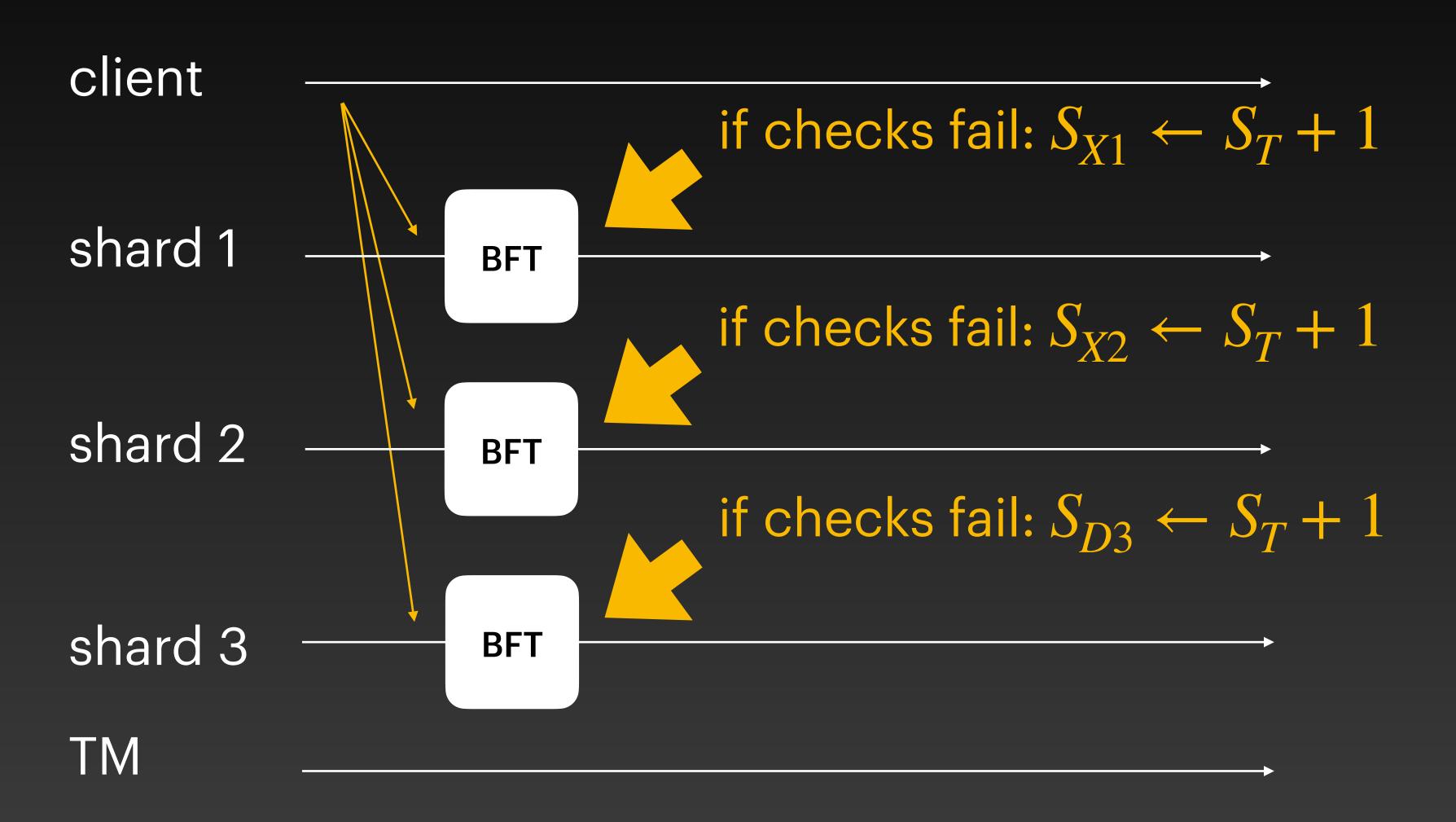
$$\{S_T, T(x_1, x_2, d_3) \rightarrow (y_1, y_2, y_3)\}$$



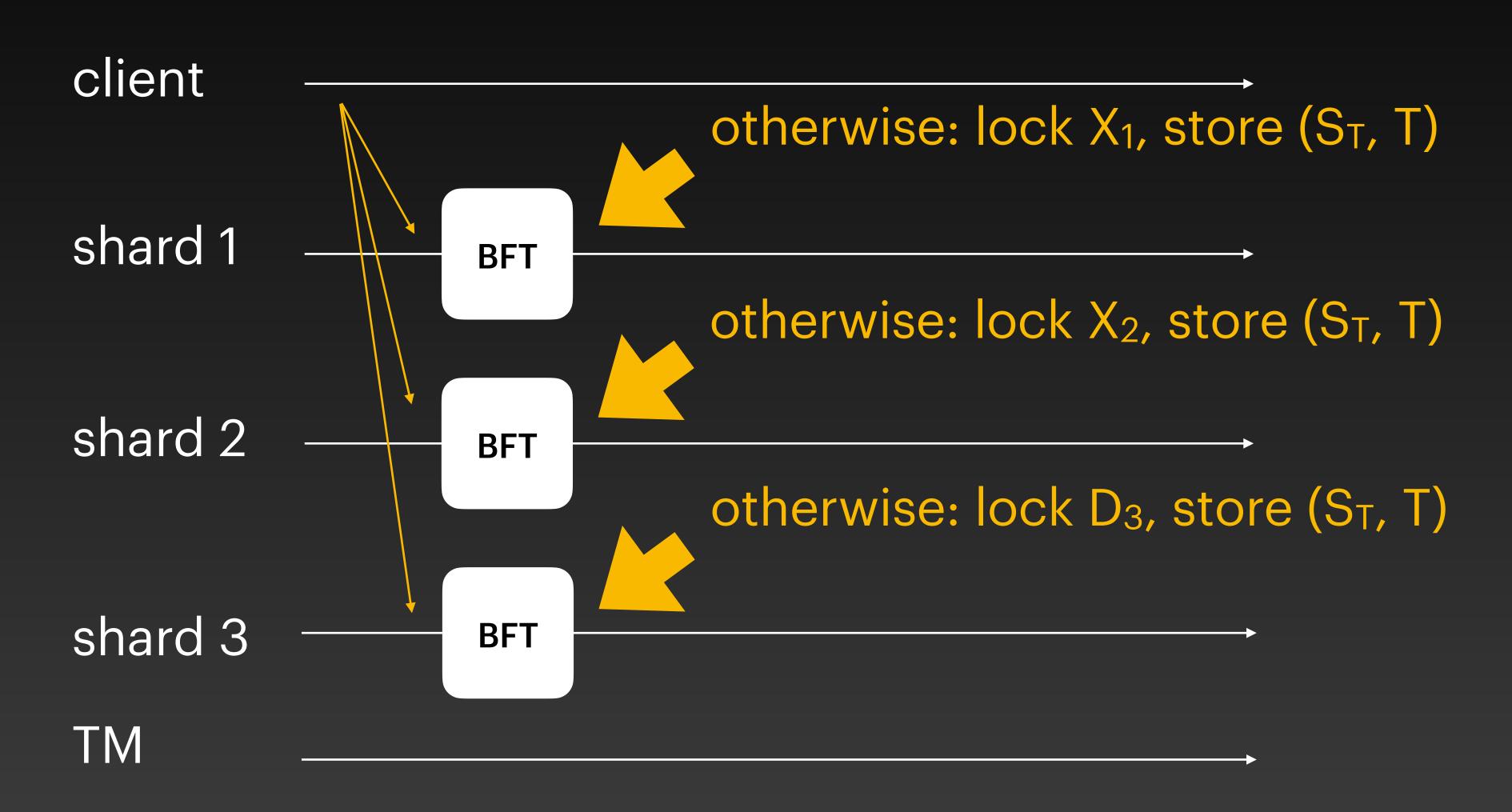
$$\{S_T, T(x_1, x_2, d_3) \rightarrow (y_1, y_2, y_3)\}$$



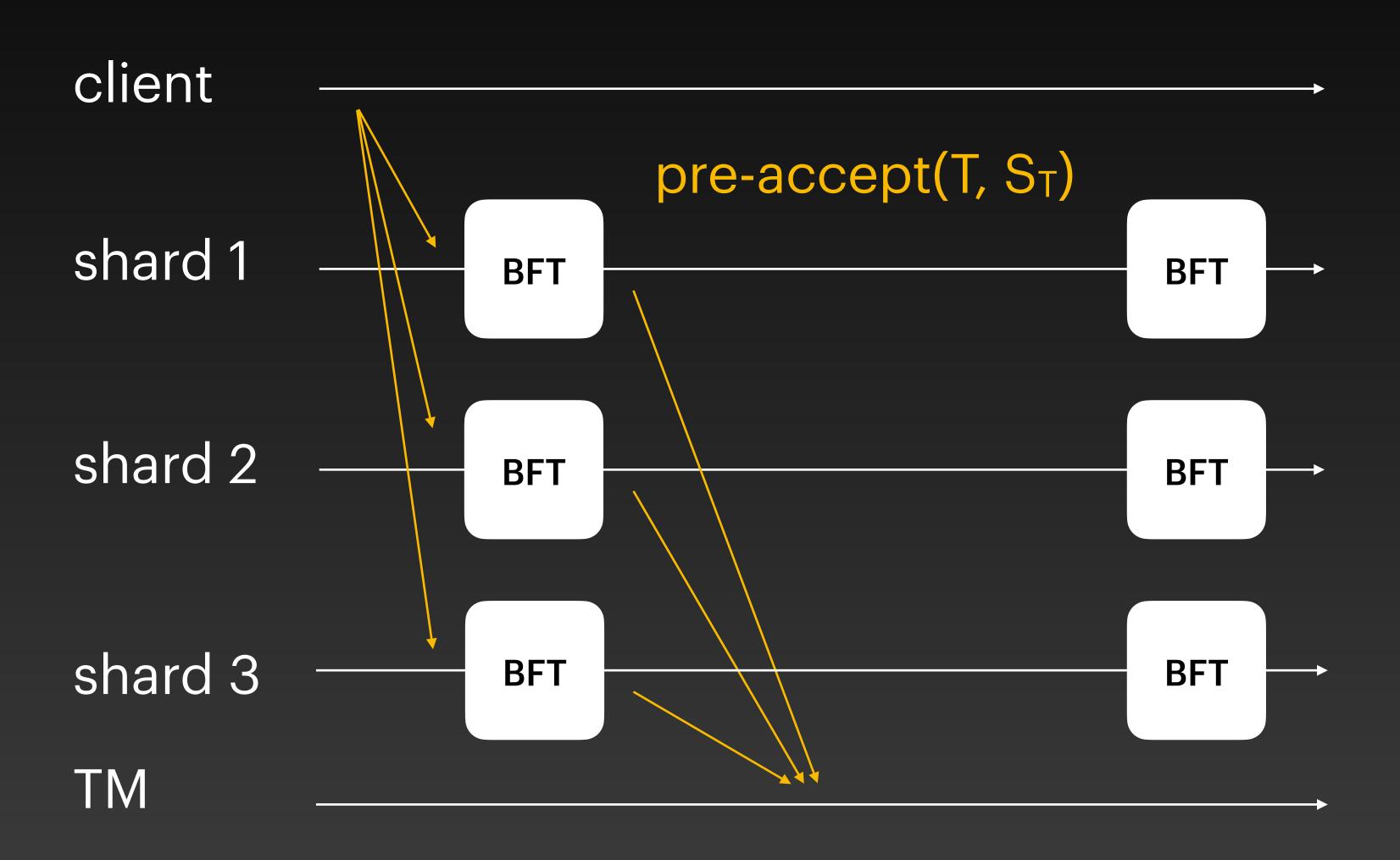
$$\{S_T, T(x_1, x_2, d_3) \rightarrow (y_1, y_2, y_3)\}$$



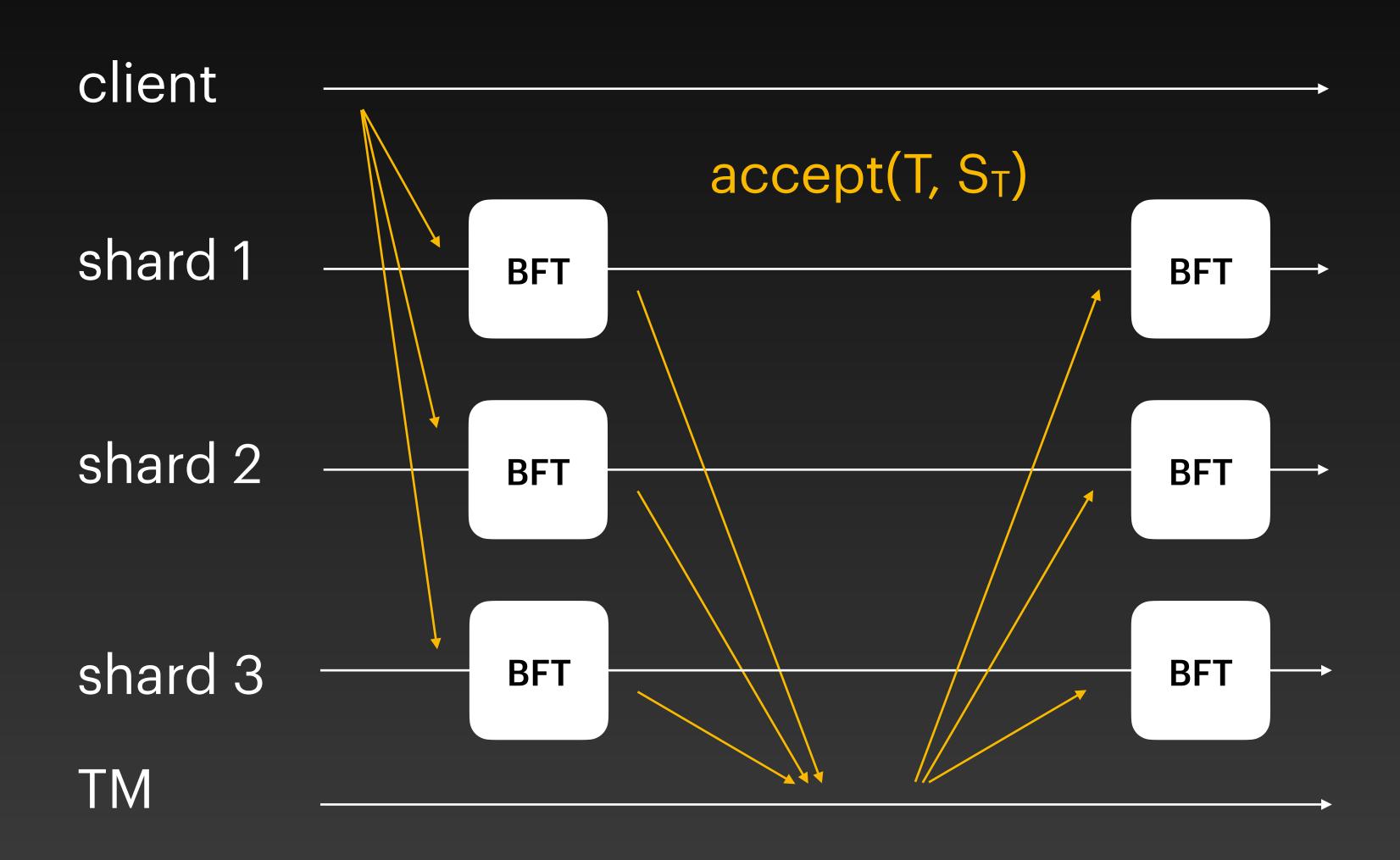
$$\{S_T, T(x_1, x_2, d_3) \rightarrow (y_1, y_2, y_3)\}$$



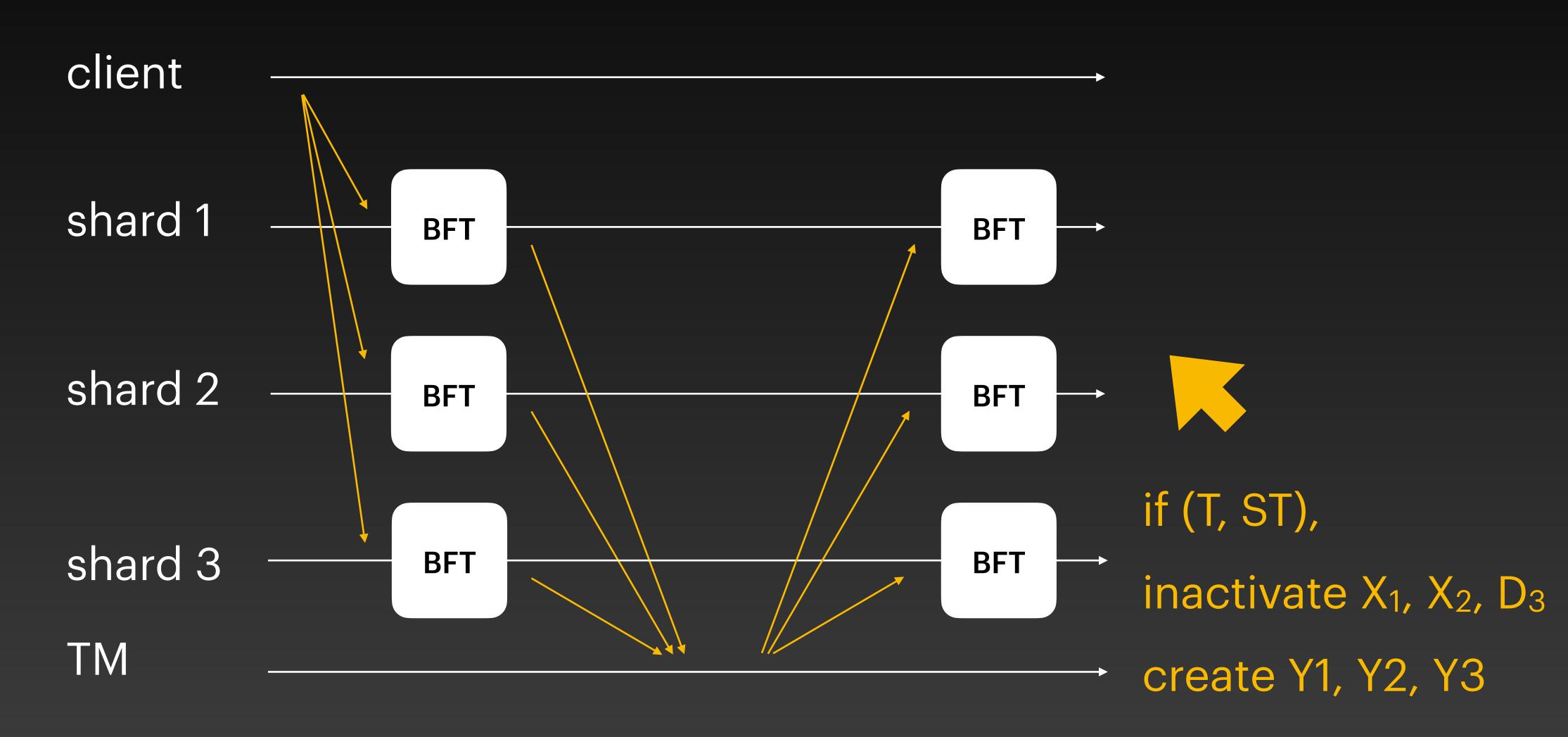
$$\{S_T, T(x_1, x_2, d_3) \rightarrow (y_1, y_2, y_3)\}$$



$$\{S_T, T(x_1, x_2, d_3) \rightarrow (y_1, y_2, y_3)\}$$



$$\{S_T, T(x_1, x_2, d_3) \rightarrow (y_1, y_2, y_3)\}$$



Why is Byzcuit secure?

Issue 1. Input shards cannot associate protocol messages to a specific protocol execution.

Sequence numbers:

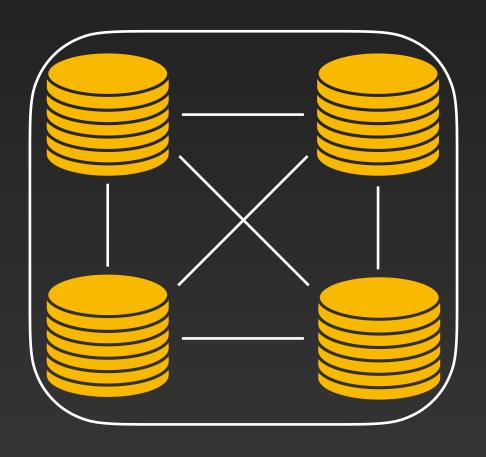
act as session ID

Issue 2. Output shards (that are not also input shards) do not experience the first phase of the protocol

Dummy objects:

all shards experience the first phase of the protocol

Anyone can be a TM



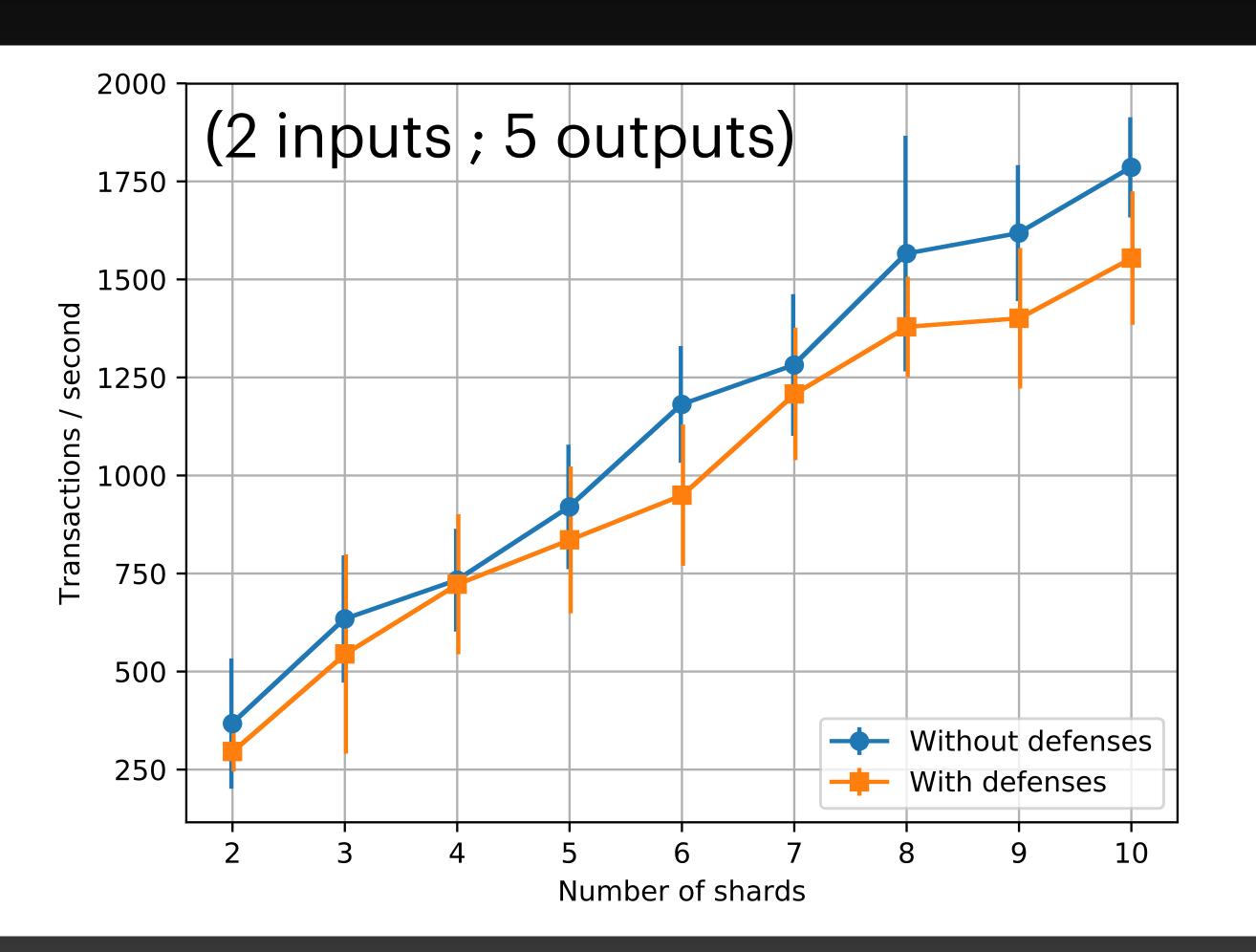


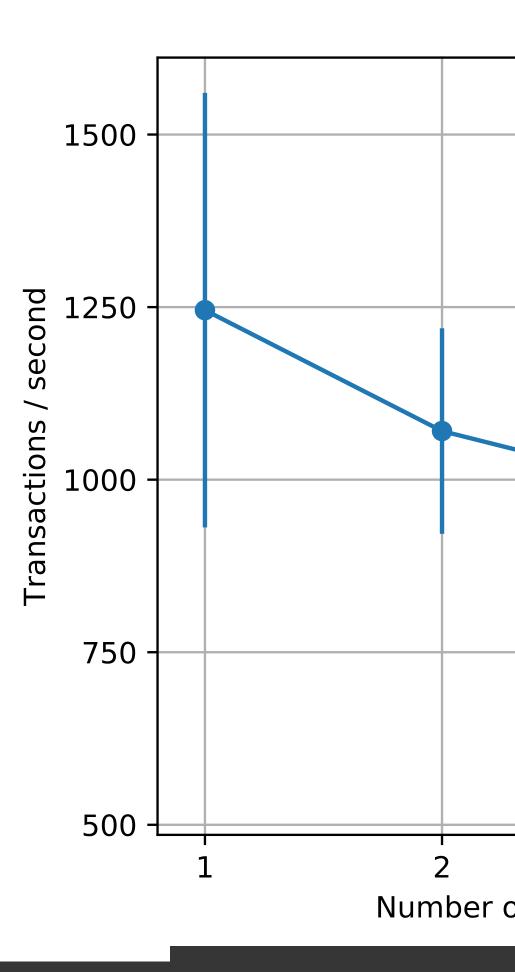
Byzcuit Implementation

- Fork of Java Chainspace
- Based on BFT-SMART
- Only a prototype to demonstrate its properties

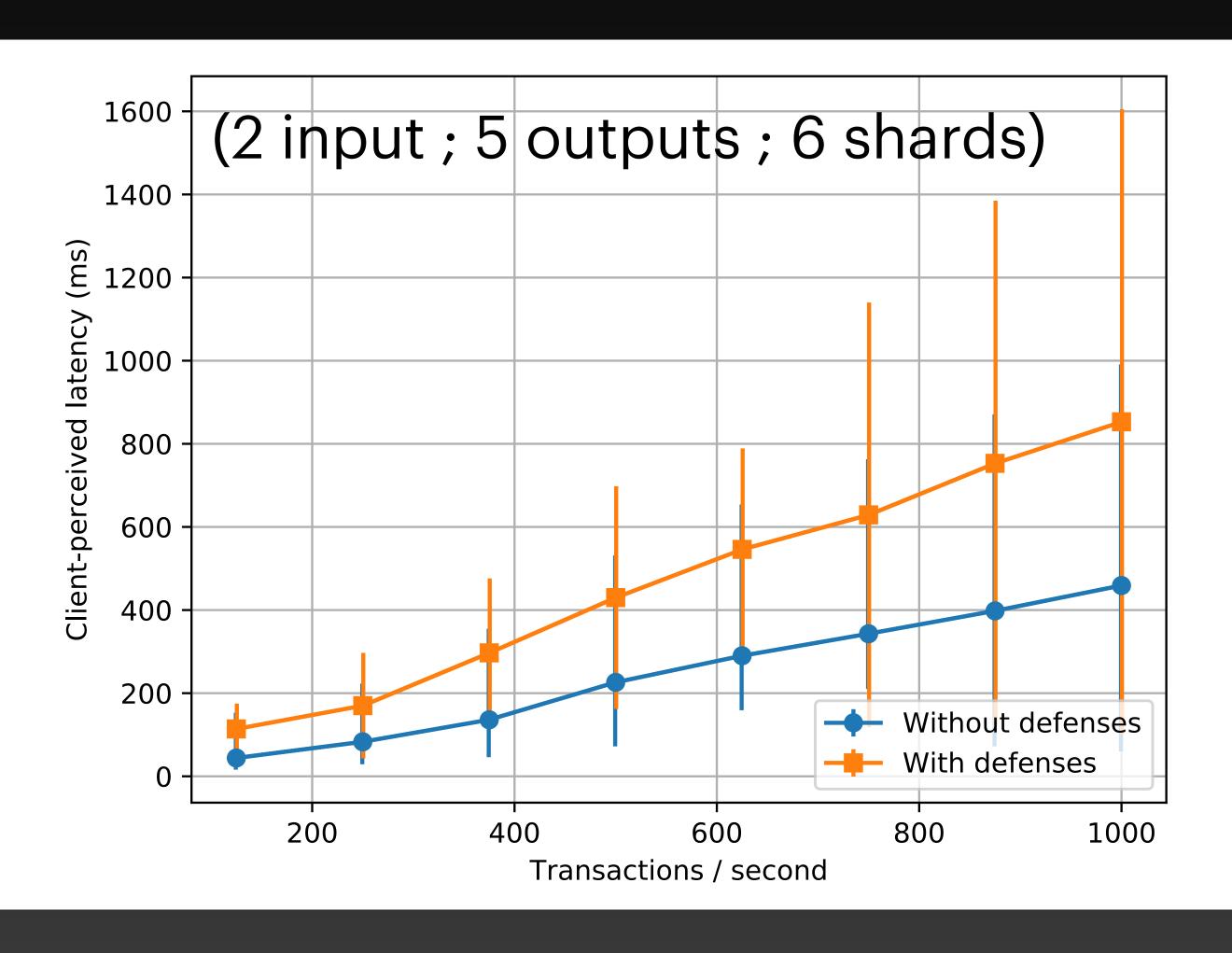
https://github.com/sheharbano/byzcuit

Byzcuit Linear scalability





Byzcuit Finality



Conclusion

Part I - Increasing Throughput

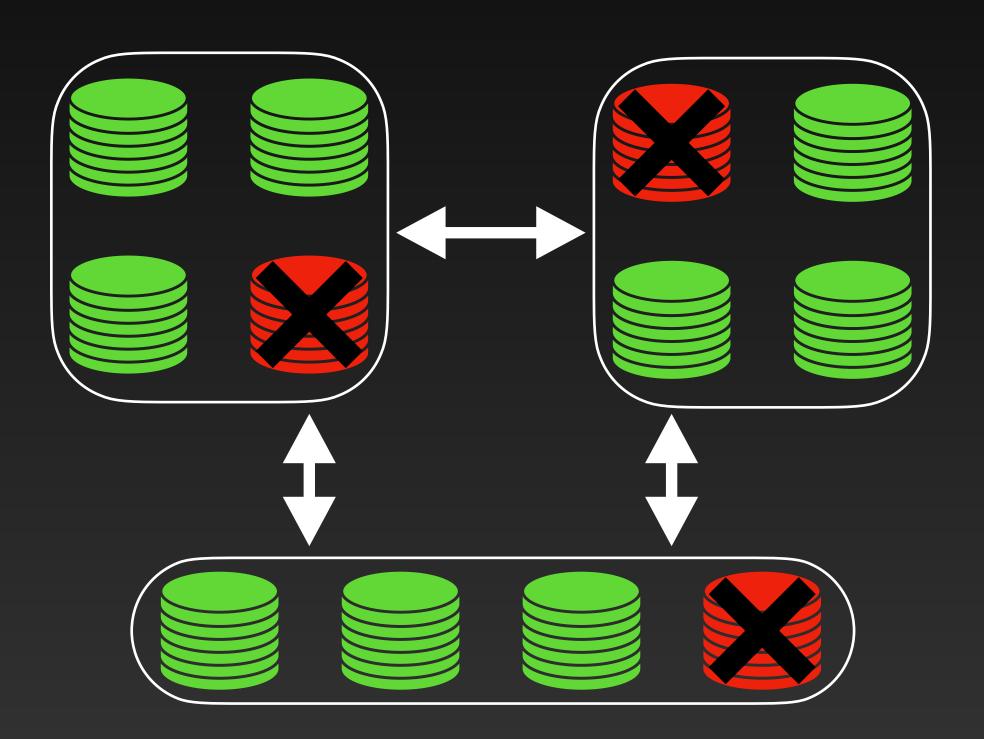
- S-BAC + Atomix
- High throughput, linear scalability, BFT resilience, Fast finality

- Paper: https://arxiv.org/abs/1901.11218
- Code: https://github.com/sheharbano/byzcuit

Reducing Latency

Using Side Infrastructures

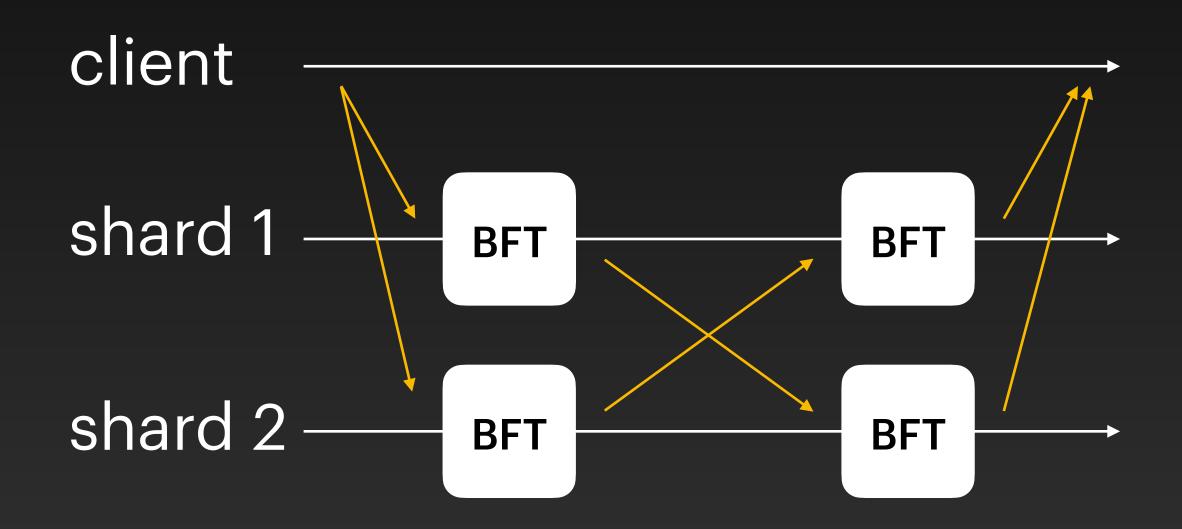
What we have so far from Part I



- Hight throughput
- Linear scalability
- BFT resilience
- Fast finality

What we have so far

and what is missing



Total Latency:

slowest shard during phase 1

+

slowest shard during phase 2

+

all communications

Make it practical for retail payment at physical points of sale

That is the ambition

What do we need? Properties

What we want

- Low latency
- BFT reliance
- Fast finality
- Hight capacity

Current industry



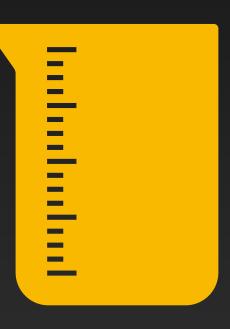
Centralized systems



Slow Finality



Little capacity



In summary

What we want

- Low latency
- BFT reliance
- Fast finality
- Hight capacity

Current industry

- Low latency
- Centralized
- Slow finality
- Little capacity

High-Performance Byzantine Fault Tolerant Settlement

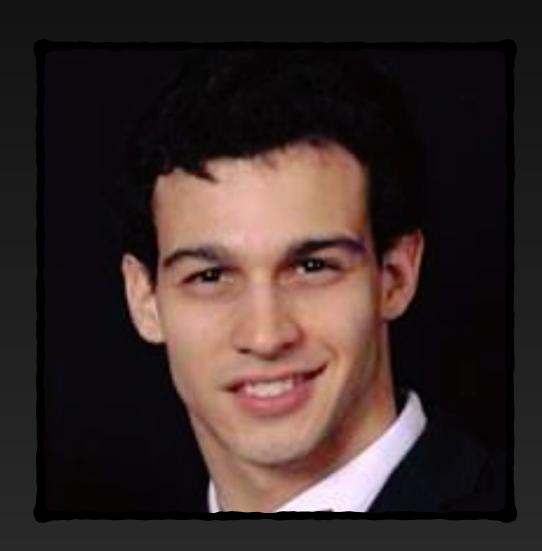
FastPay Acknowledgments



Mathieu Baudet



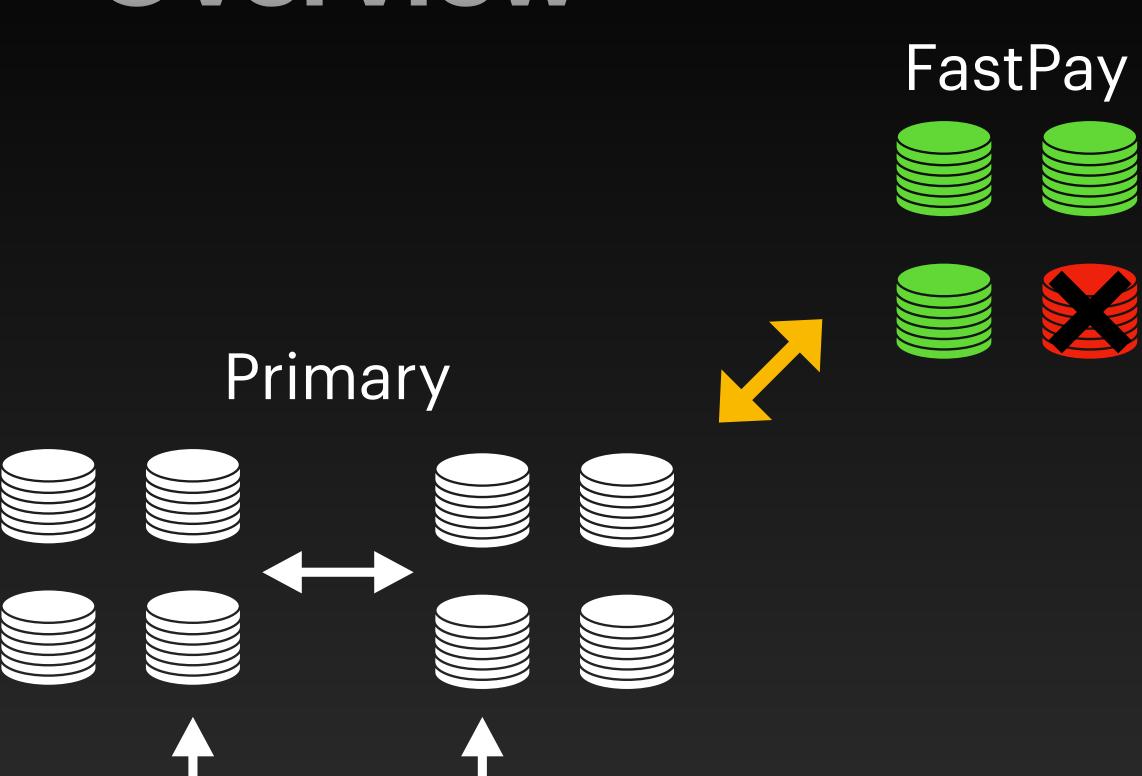
George Danezis



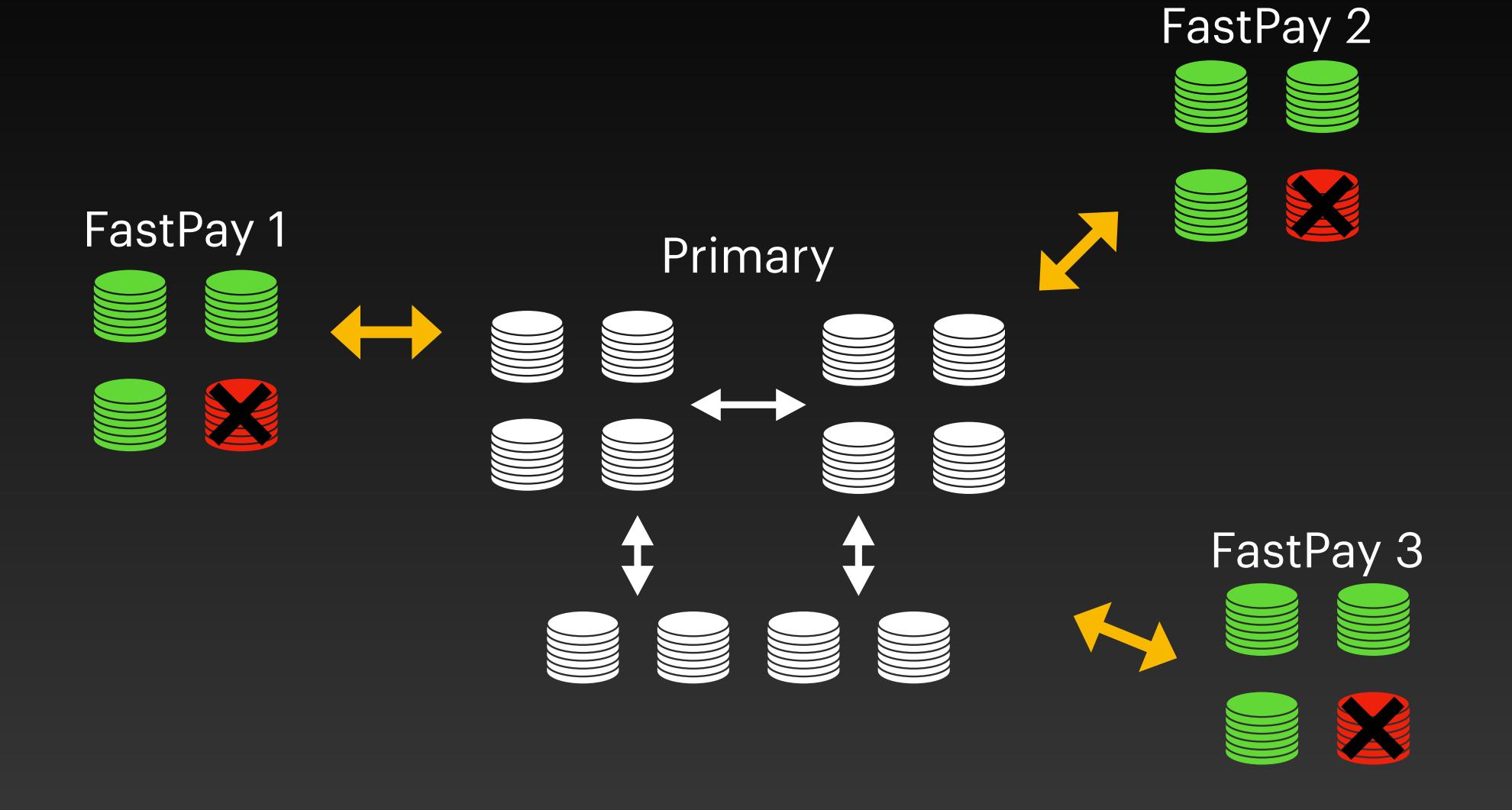
Alberto Sonnino

Facebook Novi

Overview

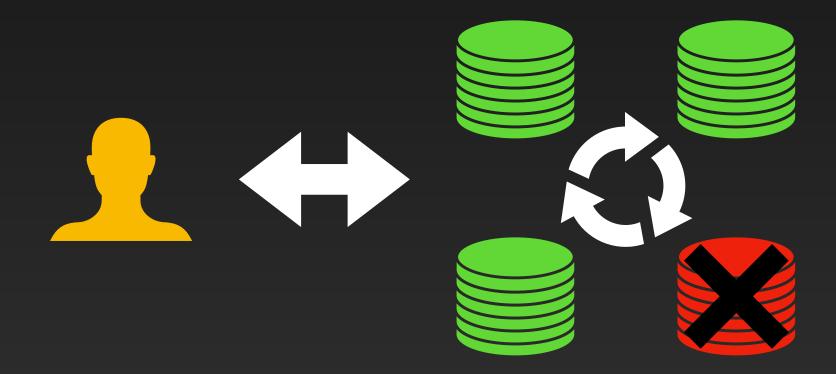


Overview



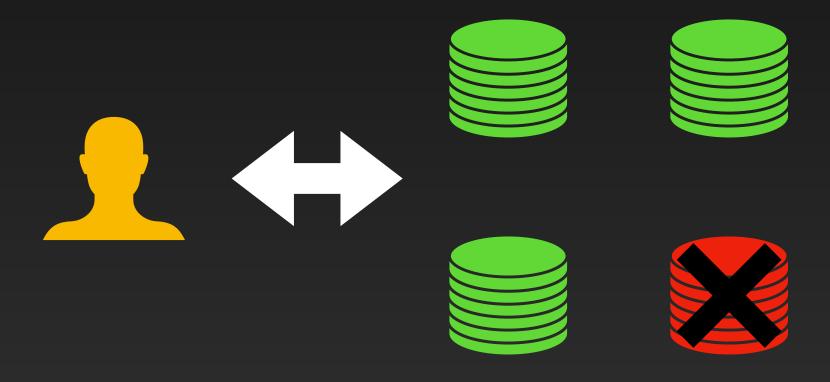
Difference with blockchains

Blockchains



Byzantine Consensus

FastPay



Byzantine Consistent Broadcast







1. transfer order





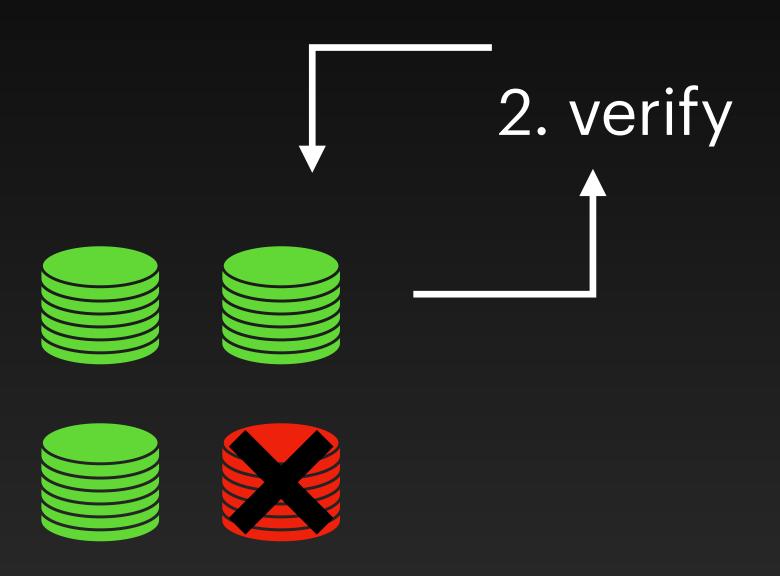




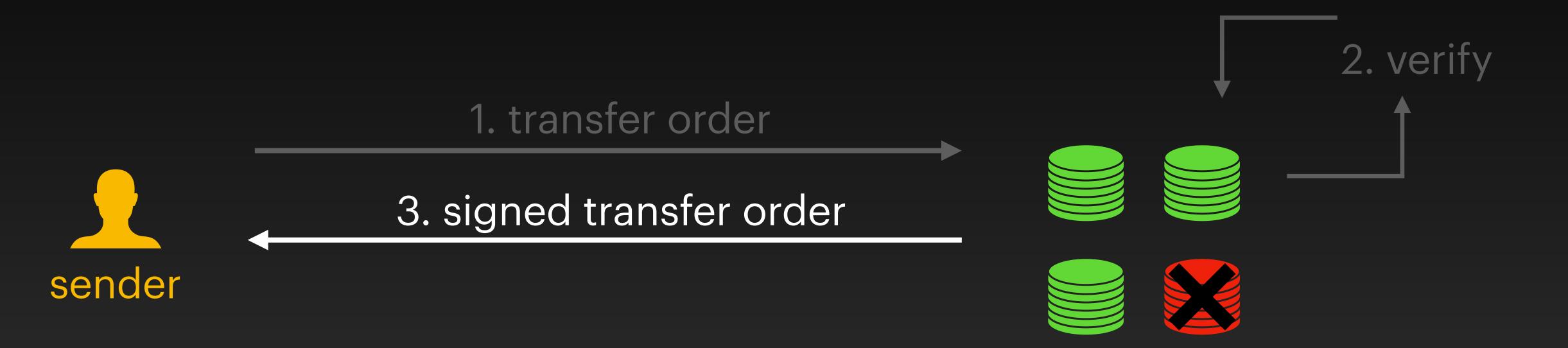


1. transfer order

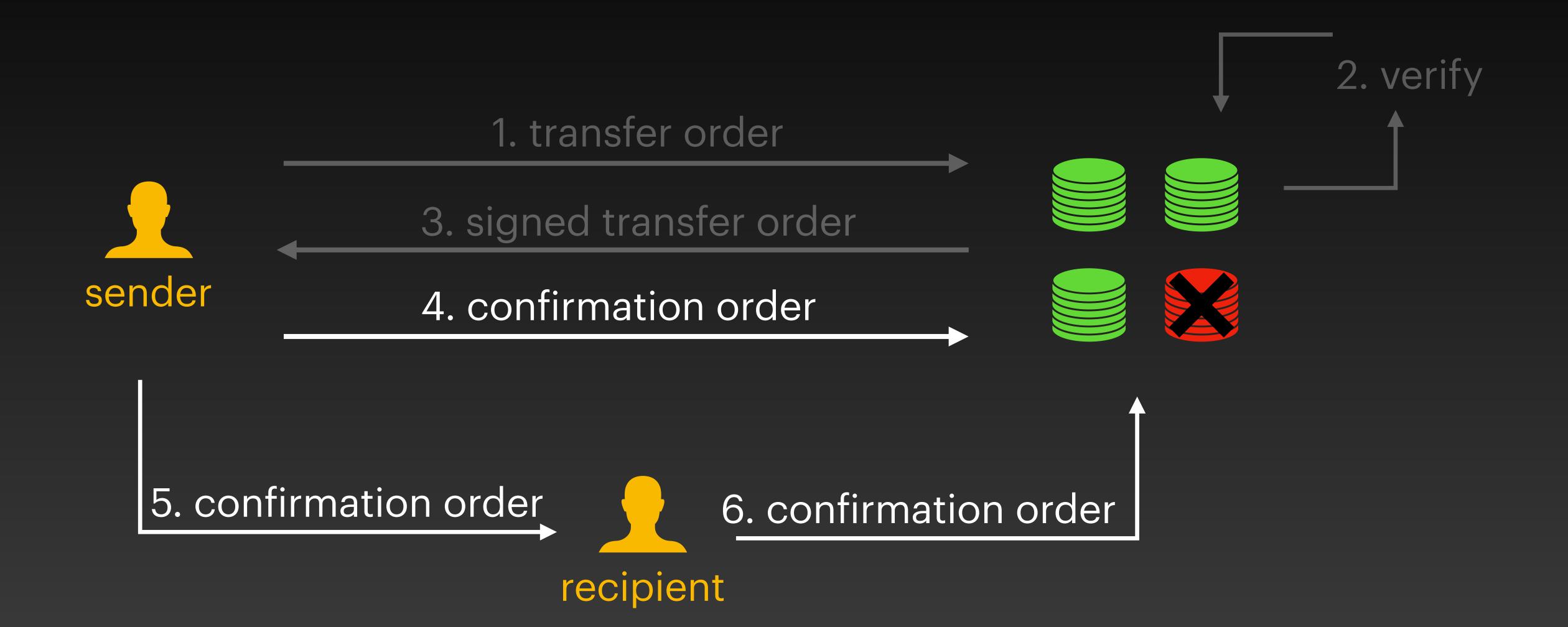


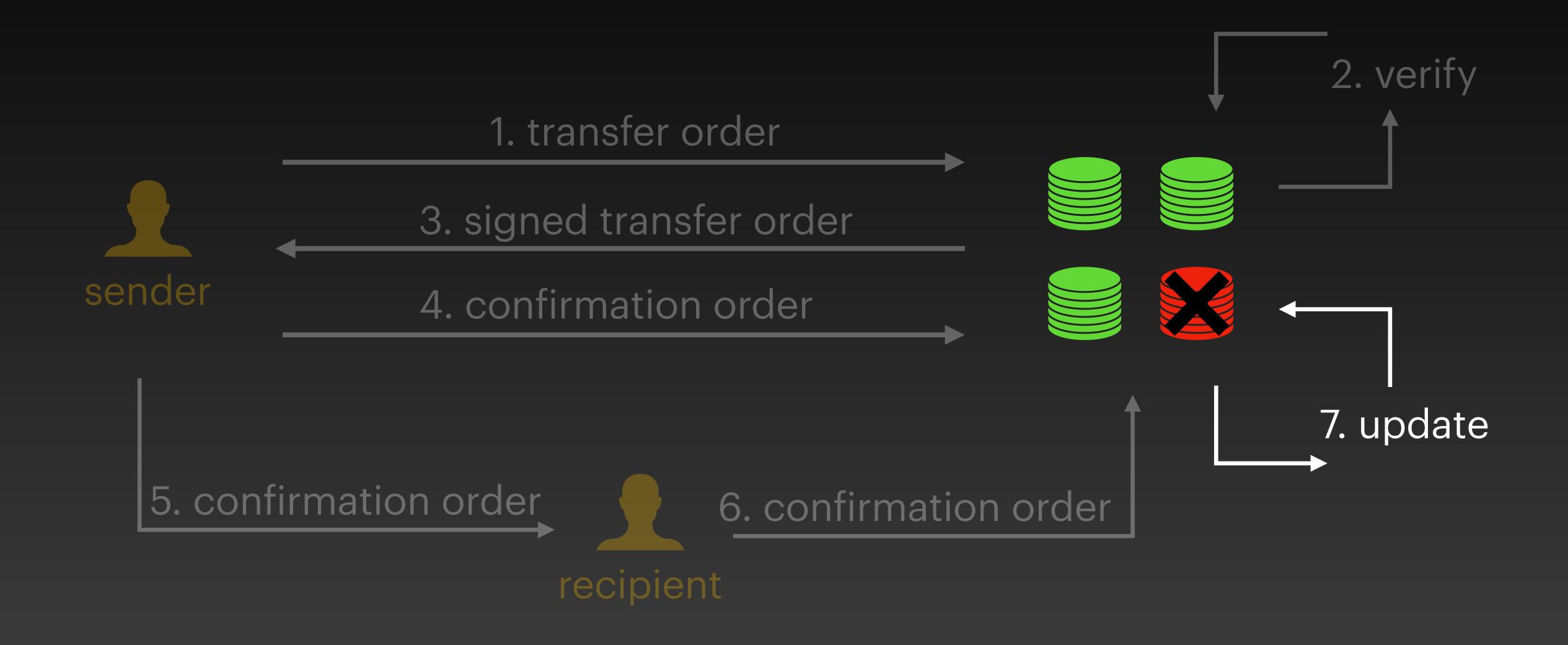




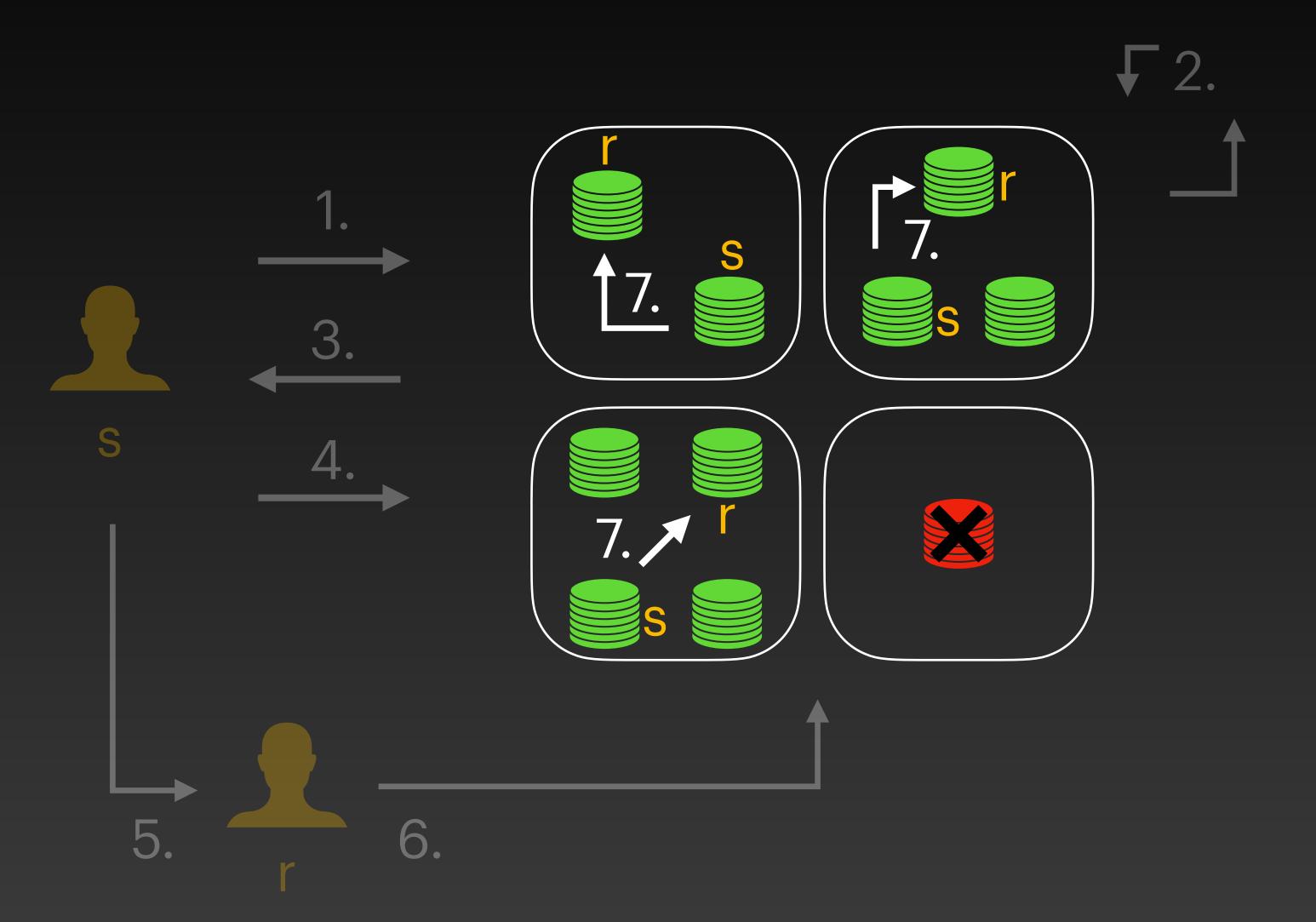




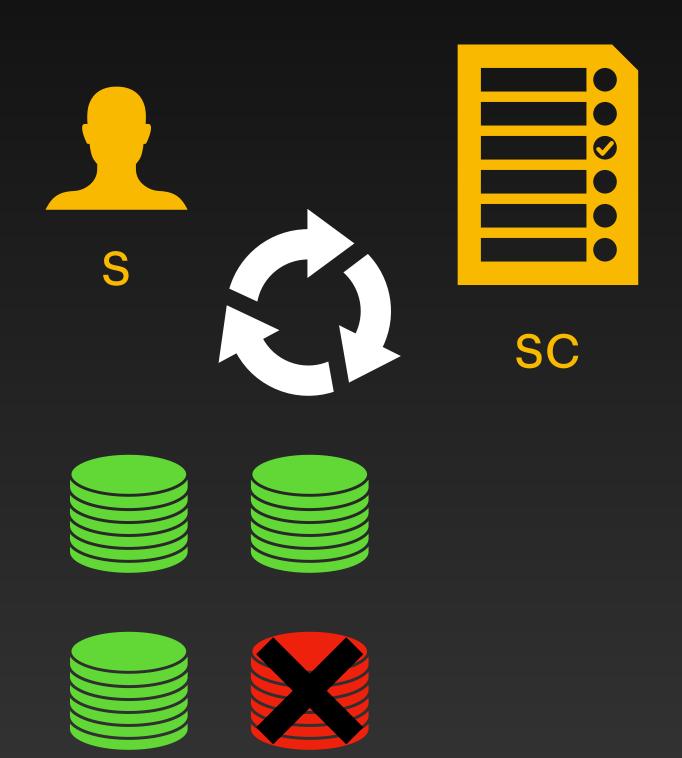




FastPay Increasing capacity



Interface it with a primary infrastructure



Smart Contract's state

- The committee information
- Total funds in the contract
- Last primary tx index
- "Redeem log"



From primary infrastructure to FastPay



1. funding transaction







From primary infrastructure to FastPay



1. funding transaction



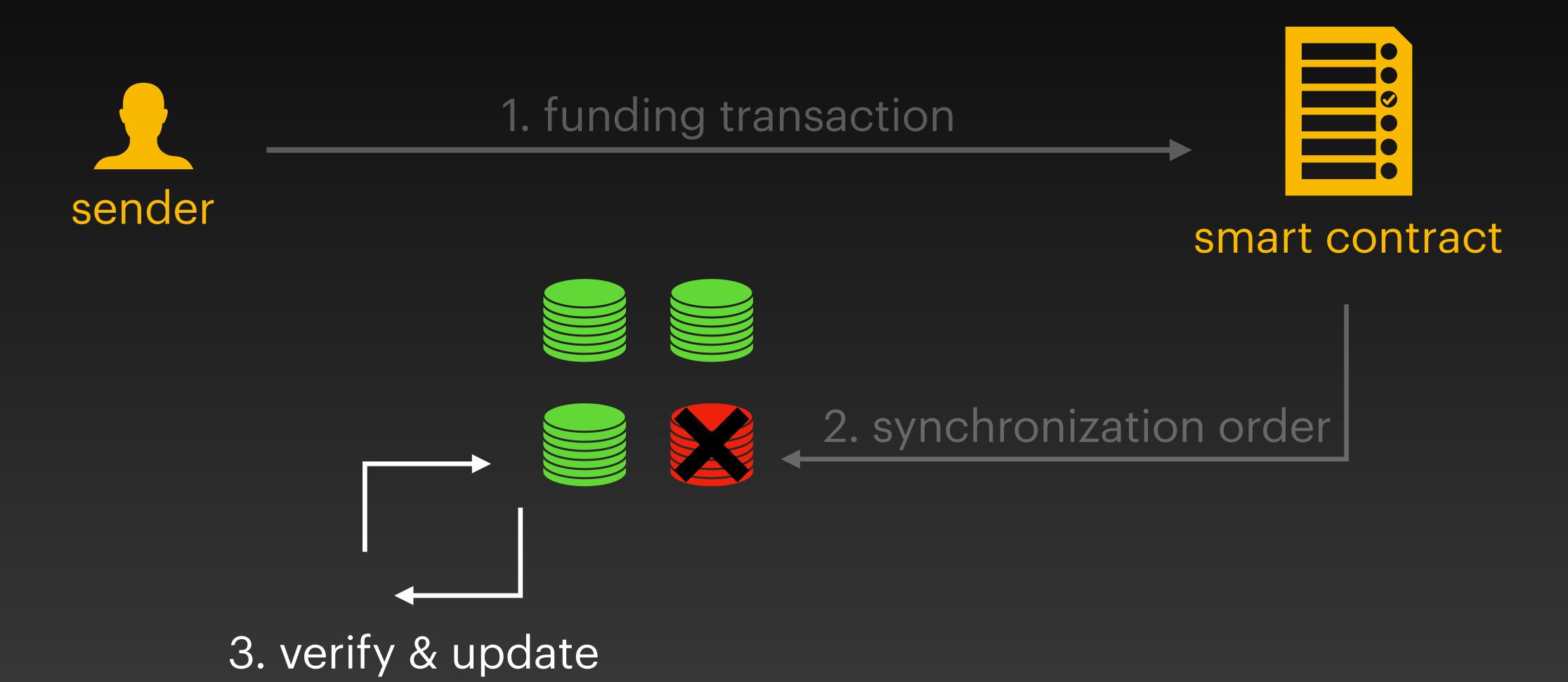
smart contract



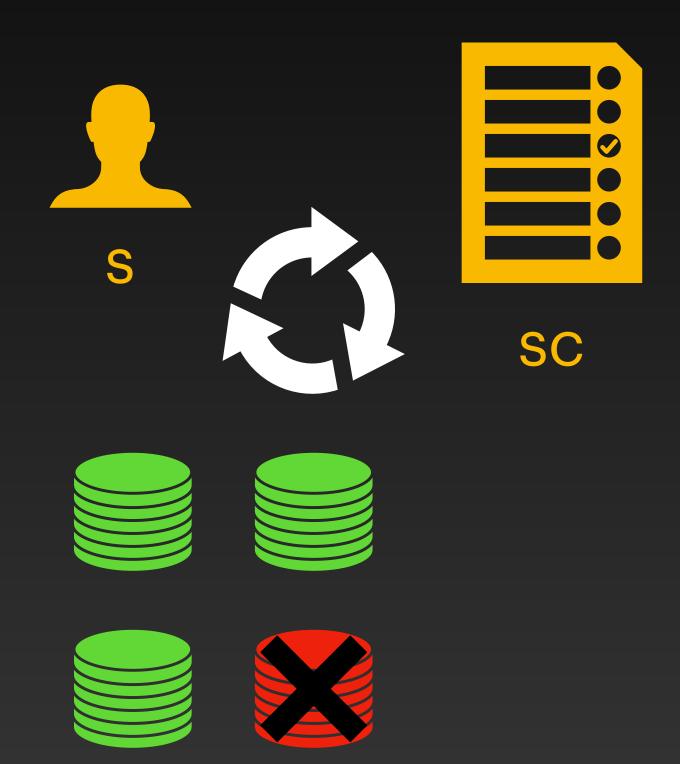




2. synchronization order



Interface it with a primary infrastructure



Smart Contract's state

- The committee information
- Total funds in the contract
- Last primary tx index
- "Redeem log"



From the primary infrastructure to FastPay

1. transfer order

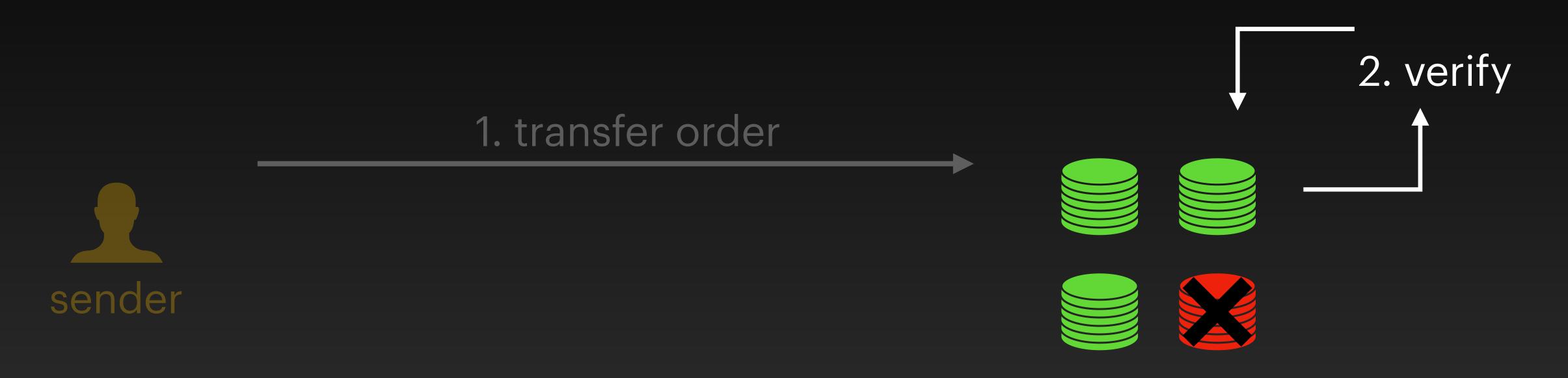




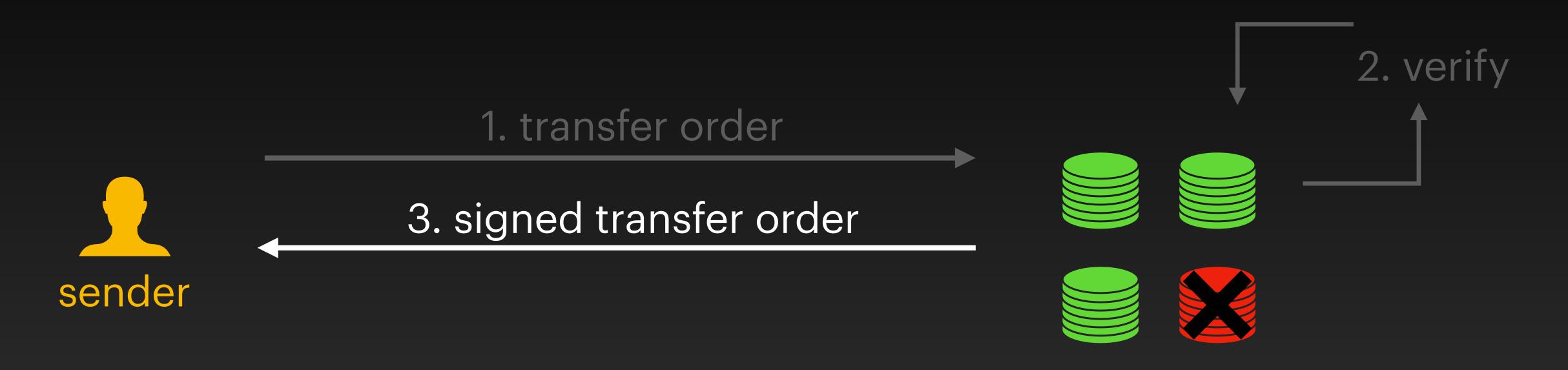




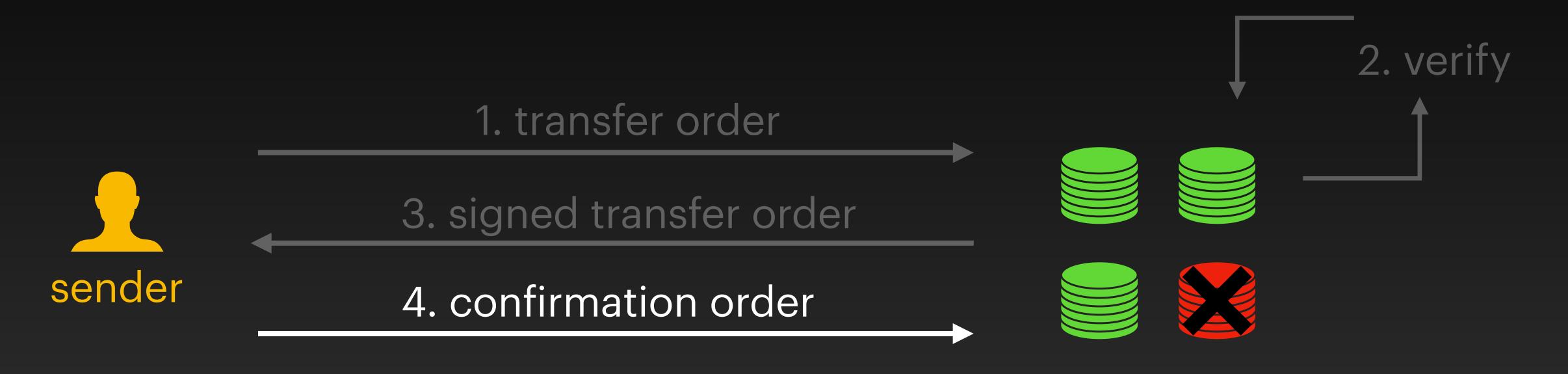




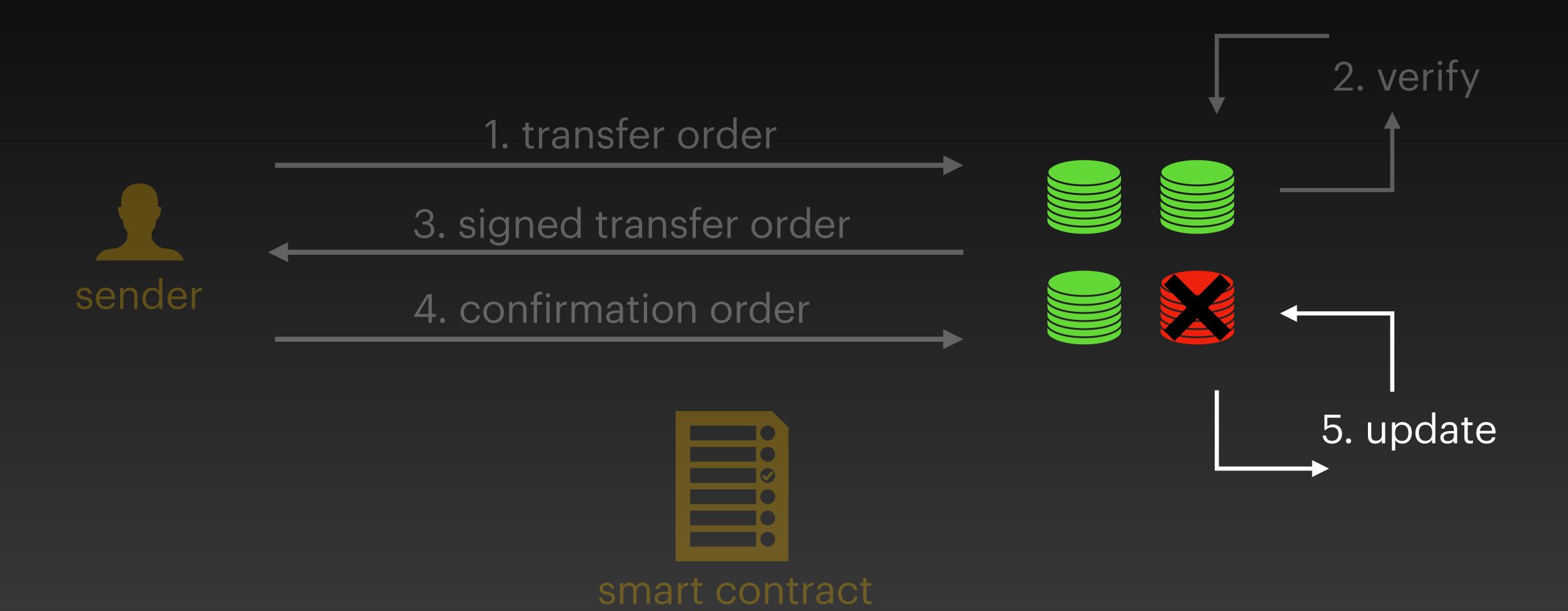


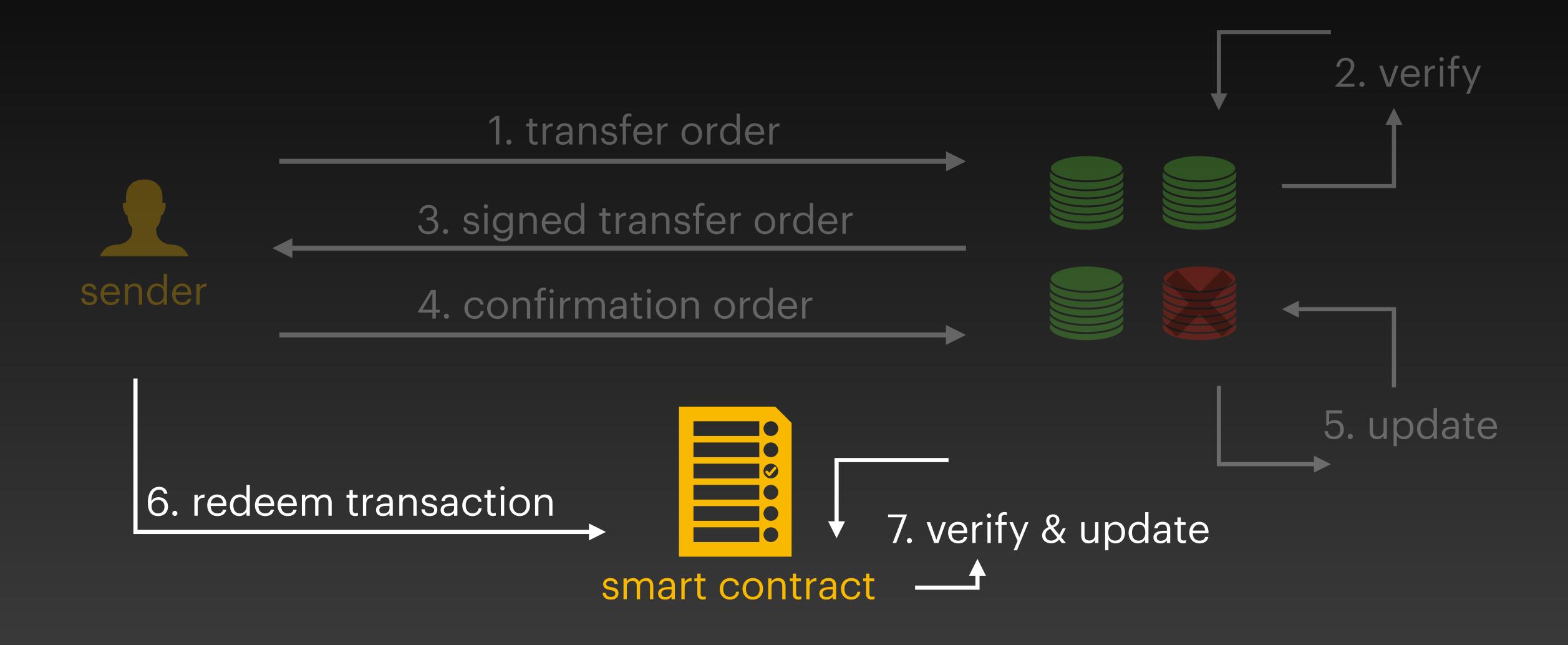












FastPay Implementation

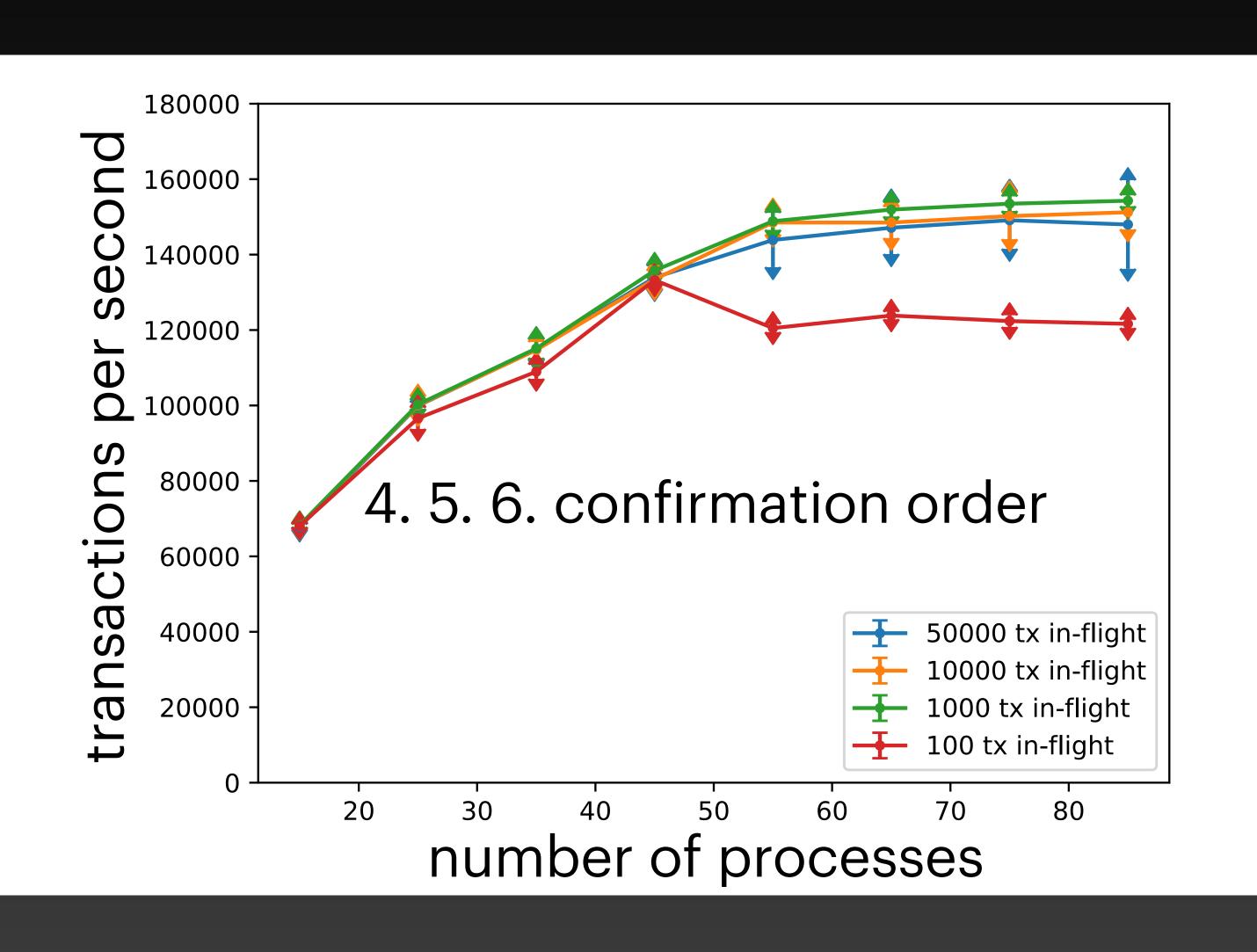
- Written in Rust
- Networking: Tokio & UDP
- Cryptography: ed25519-dalek

https://github.com/calibra/fastpay

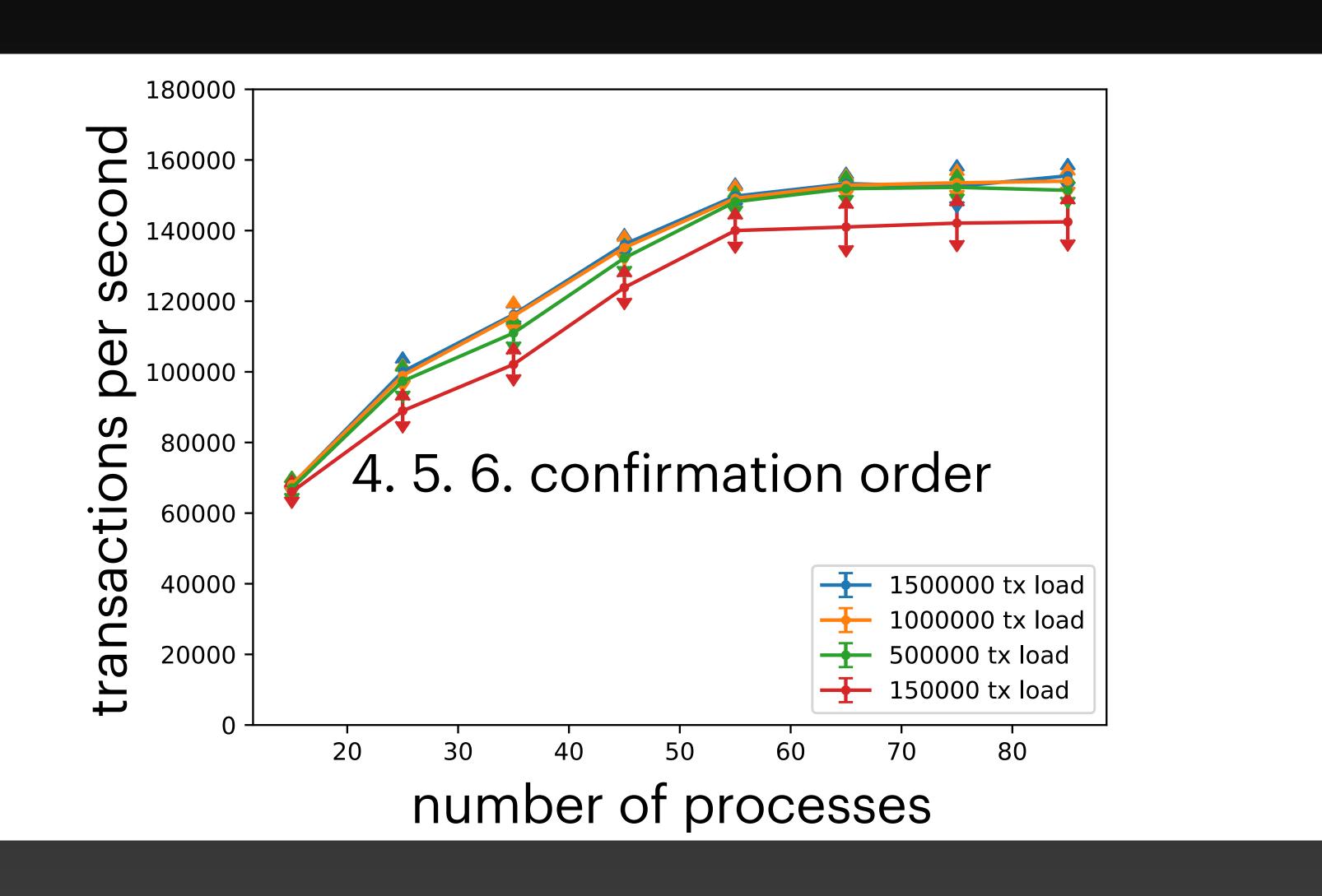
FastPay
Throughput Evaluation



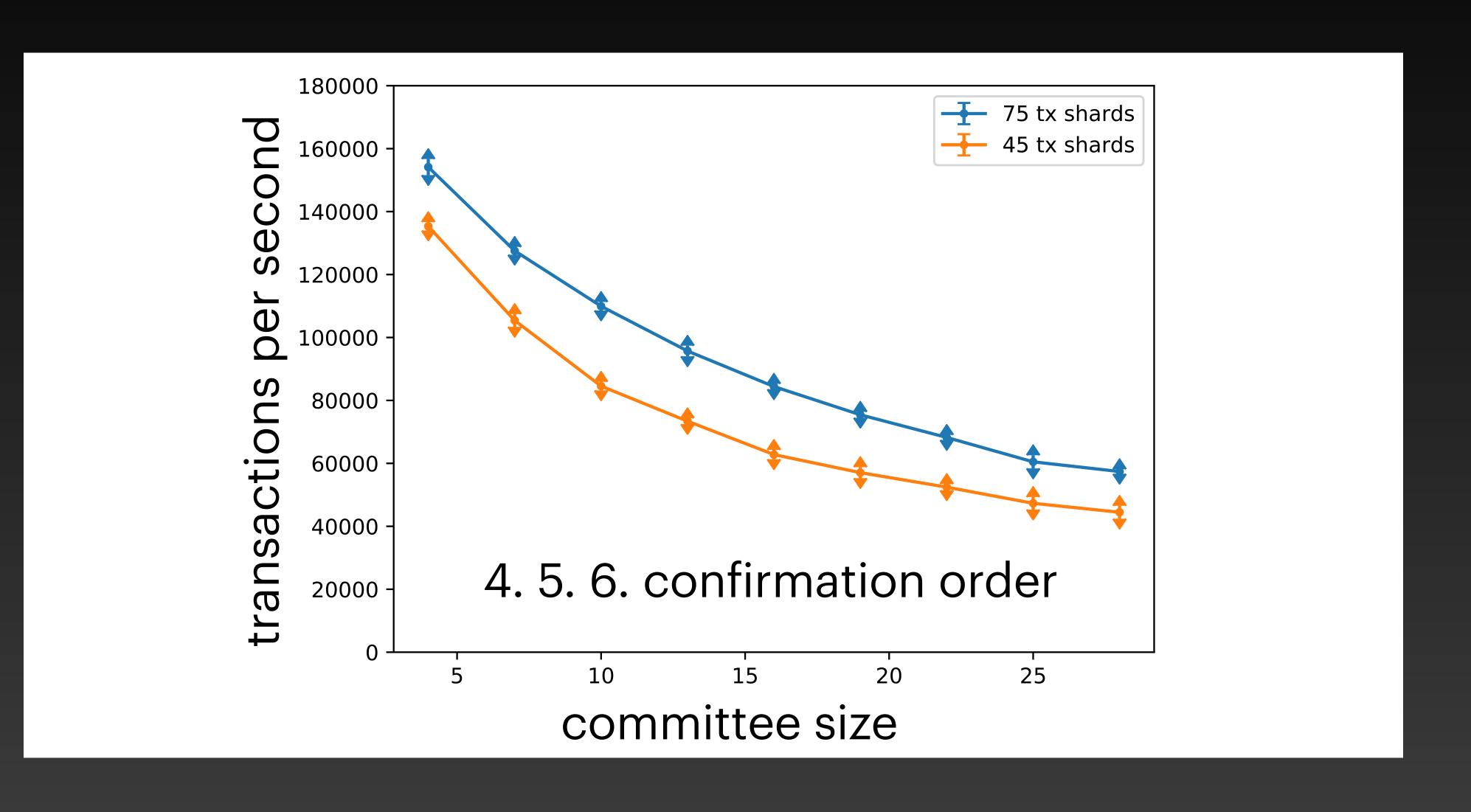
FastPay High concurrency



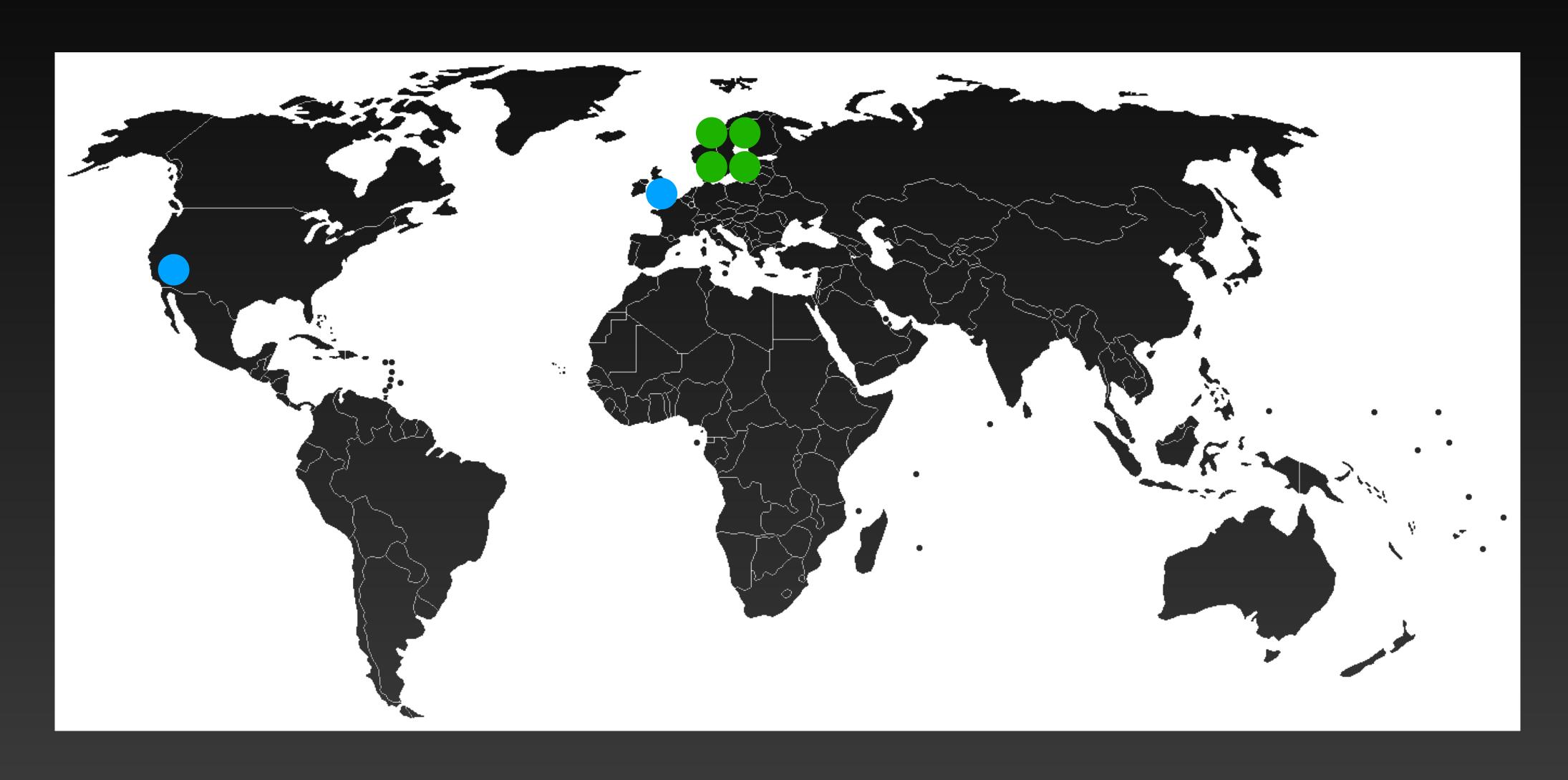
FastPay Robustness



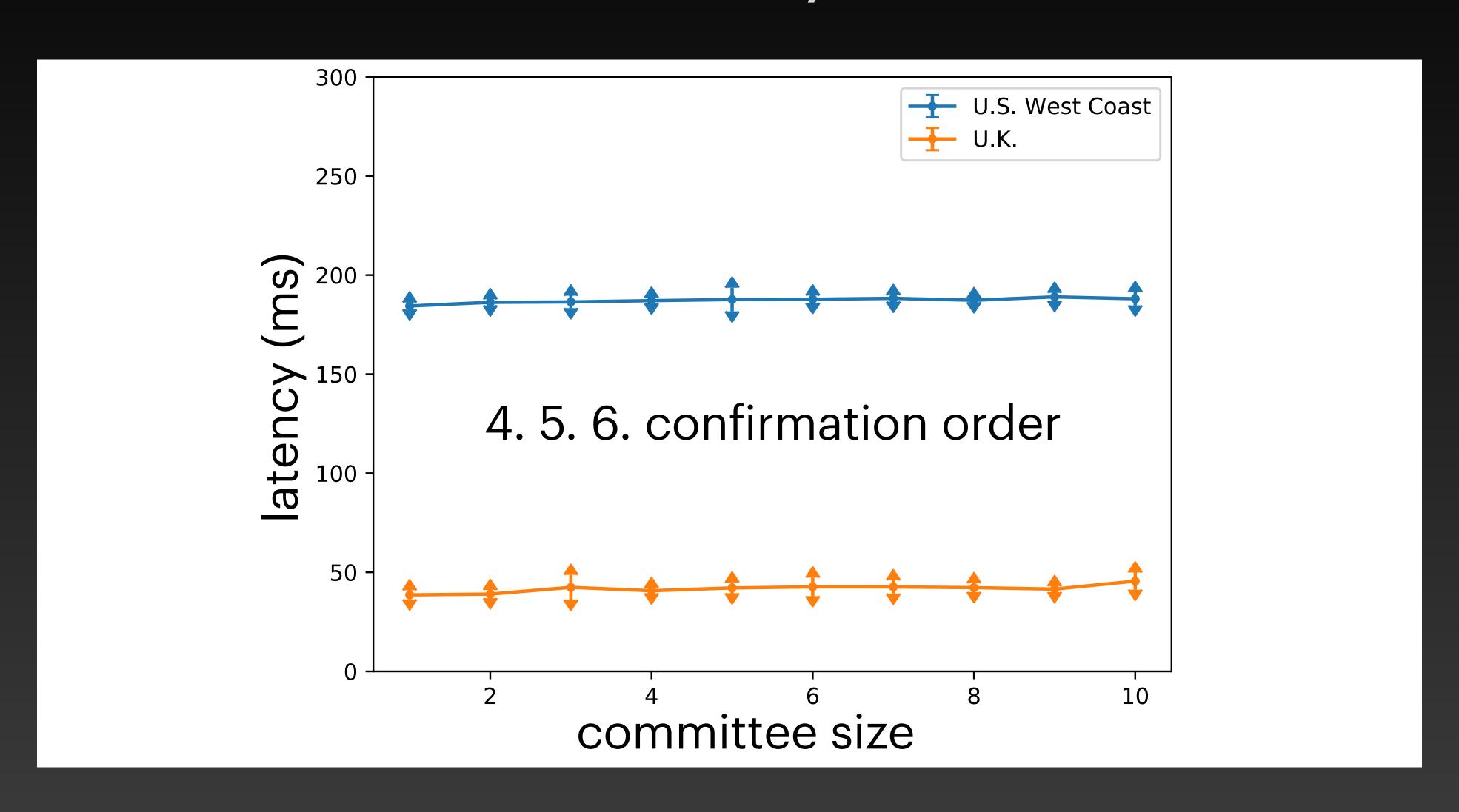
FastPay Influence of the number of authorities



FastPay Latency setup

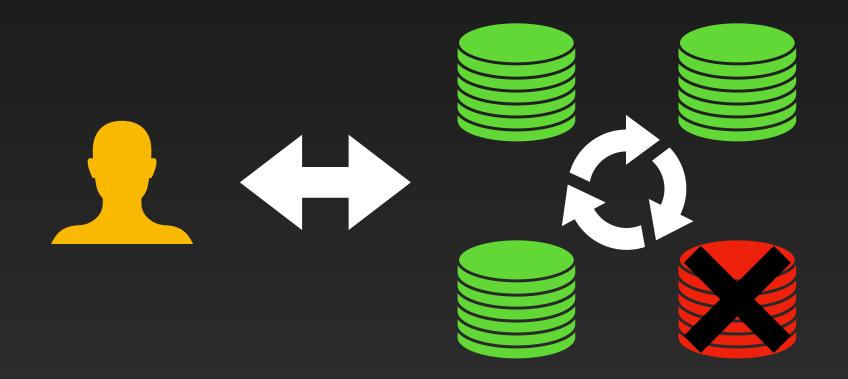


FastPay Latency



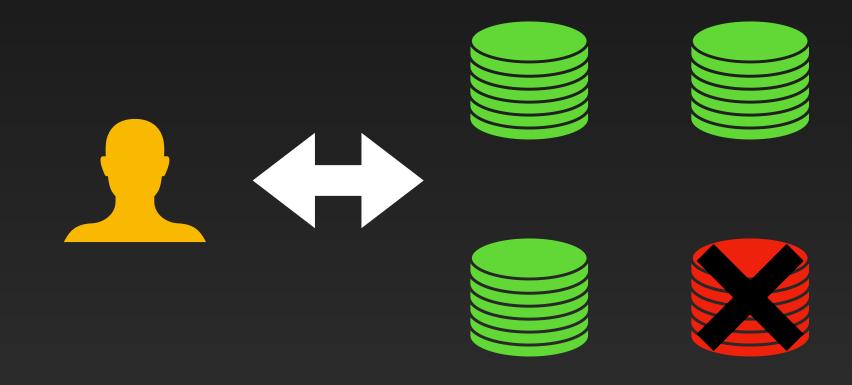
Worst-case efficiency

Blockchains



Bad leader can slow down the protocol

FastPay



No leader, nothing changes

Simplicity favors robustness and performance

Lesson I

Making it simple is hard

Lesson II

FastPay The cost of simplicity

- Less than 4,000 LOC
- Over 1,500 Git commits
- Took 2.5 months to 3 engineers

FastPay Deployment costs

- AWS m5d.8xlarge instance
- ~ 5 USD / hour

Conclusion Part II - Reducing Latency

FastPay

- Based on Byzantine Consistent Broadcast
- Simple design, low latency, high capacity, very robust

- Paper: https://arxiv.org/abs/2003.11506
- Code: https://github.com/calibra/fastpay

Scaling Blockchains

Achieving Arbitrary Throughput and Sub-Second Latency

Outline

Part I: Increase throughput

Part II:
Reduce latency

Through sharding

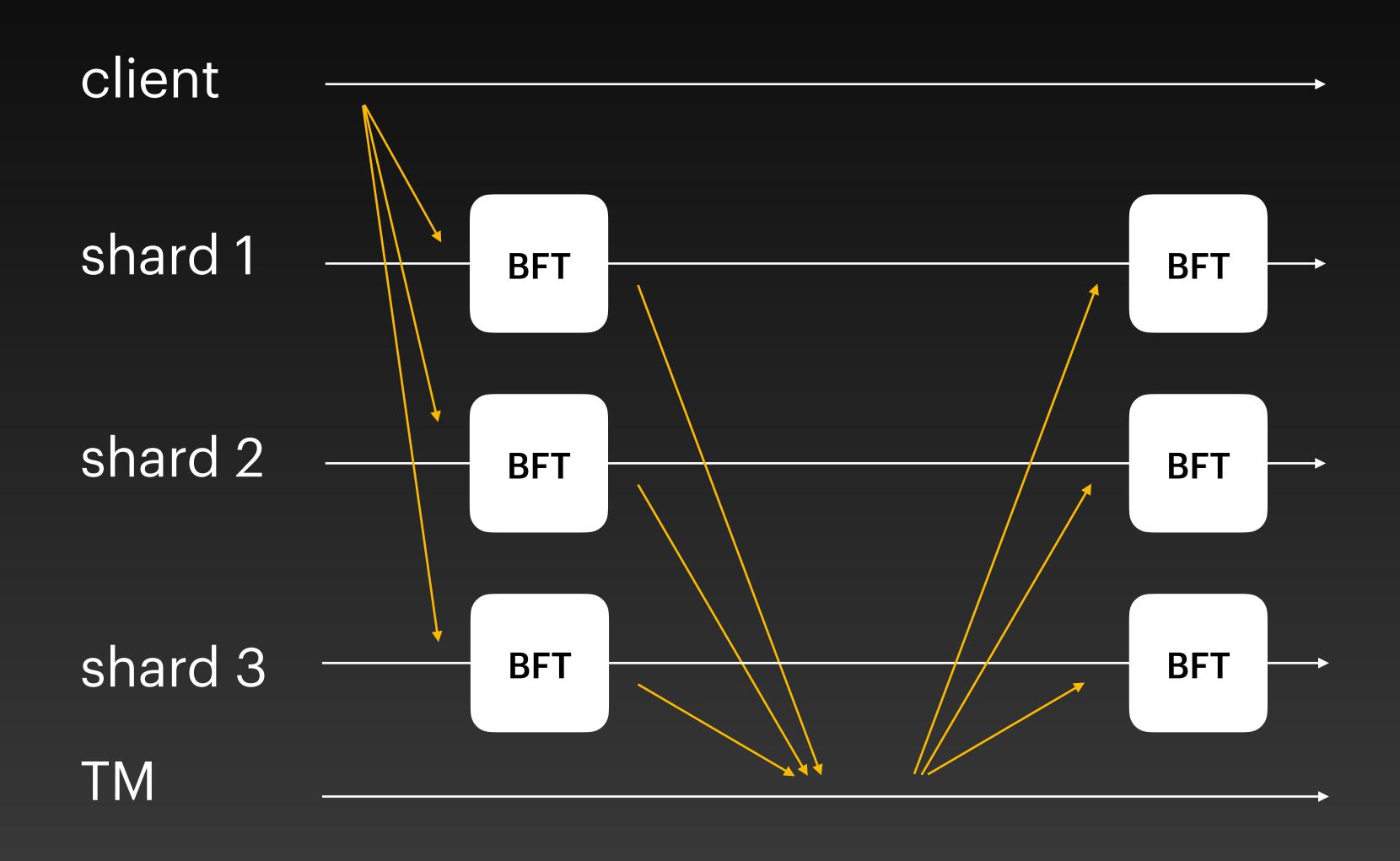
with side infrastructures

FastPay 2 Part Part FastPay 1 Byzcuit Part FastPay 3

Part

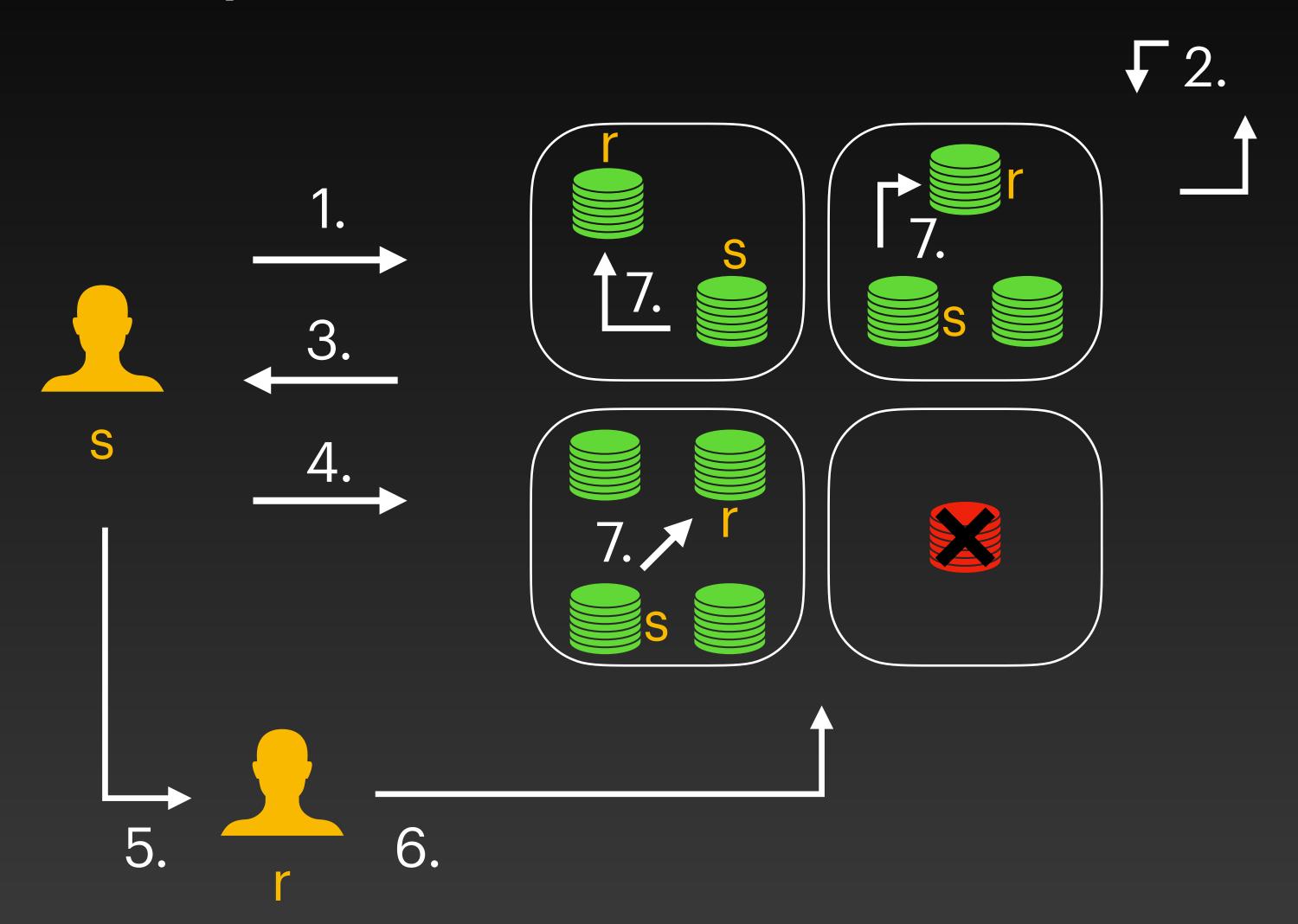
Cross-Shard Consensus

Tricky to implement right



BFT Side-Infrastructures

Byzantine Consistent Broadcast



Main Takeaways

Part I: Increase throughput

Key concepts of sharded distributed ledgers

Main challenges in building secure sharded systems in practice

Part II: Reduce latency

Side-infrastructures to bring blockchain-based payment systems to physical points of sales

How to integrate those infrastructures into a primary distributed ledger

asonnino@fb.com

Facebook Novi