SoK: Not Quite Water Under the Bridge: Review of Cross-Chain Bridge Hacks

Sung-Shine Lee  
Quantstamp, Inc.  
martinet@quantstamp.com

Alexandr Murashkin  
Quantstamp, Inc.  
alex@quantstamp.com

Martin Derka  
Quantstamp, Inc.  
martin@quantstamp.com

Jan Gorzny  
Quantstamp, Inc.  
jan@quantstamp.com

Abstract—The blockchain ecosystem has evolved into a multi-chain world with various blockchains vying for use. Although each blockchain may have its own native cryptocurrency or digital assets, there are use cases to transfer these assets between blockchains. Systems that bring these digital assets across blockchains are called bridges, and have become important parts of the ecosystem. The designs of bridges vary and range from quite primitive to extremely complex. However, they typically consist of smart contracts holding and releasing digital assets, as well as nodes that help facilitate user interactions between chains. In this paper we first provide a high level breakdown of components in a bridge and the different processes for some bridge designs. Then, we analyse past exploits in the blockchain ecosystem that specifically targeted bridges. In doing this, we identify risks associated with bridge components.

I. INTRODUCTION

In recent years, the blockchain ecosystem has evolved into a multi-chain world. Various blockchains, like the popular Bitcoin Network [1] or Ethereum [2], are evolving simultaneously. These blockchains often have their own native cryptography-based digital asset or cryptocurrency, like Bitcoin or Ether. Advanced blockchains like Ethereum support automatically executed pieces of code, so-called smart contracts, which enable programs to be developed on these blockchains. In turn, these programs often introduce additional digital assets, like non-fungible tokens (NFTs).

However, for various reasons it is often desirable to move digital assets from one blockchain to another. For example, a user may wish to move their Bitcoin onto Ethereum to deposit it into various Decentralized Finance (DeFi) protocols which may allow the user to earn interest on their Bitcoin, akin to a savings account from a bank (see e.g., [3]). In another situation, a user may have an Ethereum-based NFT which can be used in games executed by smart contracts. However, the gas fees – the cost of executing a transaction on Ethereum – may be prohibitively large, and so the simulated race may be built on a so-called layer two scaling solution built on top of Ethereum [4]. Examples of these scaling solutions include rollups (also known as commit-chains [5]), Plasma [6], or a side-chain (see e.g., [7]). These scaling solutions have their security tied to Ethereum but are able to reduce the cost of gas fees, and are therefore more attractive for applications such as an NFT-based game.

The ability to “move” a digital asset from one blockchain to another requires a protocol, which can be implemented with the support of smart contracts. Such a protocol is required to ensure that an asset can only be used on one blockchain at a time, in order to prevent double spending. Double spending is when one spends a cryptocurrency token twice [8]. In this case, the double spend would be one spend of the token on its original blockchain and one spend of the token on the blockchain it was moved to. However, since a blockchain is a self-contained system, it is impossible to actually take a digital asset from one source blockchain and put it on a different destination blockchain. Instead, a representation of the original asset on the source blockchain must be created on the destination blockchain. Thus such a protocol must involve some cross-chain communication (see Figure 1), and this communication protocol (and its implementation) is called a bridge.

Bridges are complicated protocols and software projects. They can be as complicated as blockchains themselves (especially if they are a decentralized protocol), and may be required to be implemented in various languages (one for each blockchain the bridge interacts with), each with its own nuances.

Moreover, bridges need to unlock or mint digital assets on destination blockchains. Bridges are therefore responsible for distributing valuable digital assets, and
as a result, have become targets for attackers. In the last year, over $1 billion USD worth of digital assets have been stolen, incorrectly minted, or locked in these systems [9]. The result is that some users are without their digital assets, confidence in cross-chain protocols is shaken, and protocols fail to operate as promised.

In order to prevent issues like this from happening again, a deep understanding of how bridges work is required. In Section II, we review the general structure of a bridge. Next, we review various bridge exploits that have occurred on the components of a bridge. In Section III, we recount exploits that have or would have targeted parts of the bridges that hold user assets in custody. Section IV describes exploits that would have taken advantage of the part of the bridge issuing asset representations on a destination blockchain. Exploits that abuse the protocol’s cross-chain communication system are described in Section V. Finally, Section VI describes exploits that arise due to poorly defined digital assets, namely, some ERC-20 tokens. We review related work in Section VII. Finally, Section VIII concludes.

II. BRIDGE ARCHITECTURE

We now describe the high level architecture of bridges. Throughout the section, we will assume that the source blockchain for an asset is Ethereum, but any blockchain that supports smart contracts can be considered without loss of generality. We also assume that assets from Ethereum are bridged to another blockchain that supports smart contracts, which we will call “Another Chain” at times.

A bridge transfers assets from a source blockchain, where the asset is originally implemented. The digital asset is implemented either natively, as in the case of Ether (ETH) on Ethereum, or as a smart contract. Example of smart contract implemented tokens are ERC-20 tokens [10] and ERC-721 tokens [11] (i.e., NFTs). The bridge enables unlocking or creating a representation of this asset on a destination blockchain. Depending on the implementations of these chains, the implementations for locking and unlocking assets on these chains may not be symmetric; however, these details are beyond the scope of this work.

The bridge’s representation of an asset on the destination blockchain may introduce different behaviour for the asset, or it may seek to provide similar behaviour for the asset. Often, the destination blockchain representation is transferable to any party on the destination blockchain, if this is a feature of the asset on the source blockchain. In these cases, the bridge smart contracts on the destination blockchain should accept this representation from any party on that blockchain, in order to move that asset back to the source blockchain. This is necessary as otherwise users on the destination blockchain may not associate the representation with the original asset. For example, if a user is given a representation of a newly minted representation of Ether on a non-Ethereum blockchain, but that representation cannot be traded for Ether on Ethereum, users are not likely to associate the same value to it or use it in the same way.

We now describe how bridges work. First, bridges have a custodian on the source blockchain: a smart contract that locks up assets that are deposited into it. On the destination blockchain, bridges have a debt issuer that can create (or mint) digital representations of tokens for those supported by the custodian. The custodian signals (e.g., through an Ethereum event) that a digital asset was received and that the corresponding debt issuer on the destination blockchain can mint a representation of the asset. As the representation of the asset can be traded for the original asset, the representation is in fact a debt token. Since each blockchain is a closed ecosystem, a communicator reads the event emission on Ethereum to send a signal for debt issuance on the destination blockchain. The blockchain is a closed ecosystem because smart contracts are passive; they cannot actively (or regularly) read from non-transaction data, including data that only exists on other blockchains. This process is illustrated in Figure 2.

The custodian-debt issuer architecture is designed to avoid double spending of digital assets that have been sent across a bridge; it is important that bridges only mint digital representations only after receiving the true asset on the source blockchain. This prevents double spending by only having one representation of the token freely transferable at a time.

To reverse the process, a user destroys (or burns) the debt token on the destination chain. The communicator observes the destination chain, looking for every event corresponding to a burn. When the burn is complete, the communicator signals the custodian that the asset can now be released on the source blockchain. This process is illustrated in Figure 3.

Blockchain systems that write data to the blockchain from external sources are called oracles (see Figure 4). An oracle is an agent that fetches external infor-
should wait until the transaction depositing assets into the custodian has a sufficient number of confirmations — blocks appended to a chain containing the deposit transaction — following it on the source chain (or on the destination chain, if signalling to the custodian to release assets instead).

Other considerations are also important and may impact the bridge’s implementation. Depending on the source and destination blockchains, a user may not have a one-to-one mappings of identities on both chains; in this case, the bridge cannot simply mint the representation of the asset to the same address on the destination chain. Moreover a digital asset may have different representations, each with a different implementation, on each different blockchain.

Not all blockchains have the same address format. Bitcoin addresses are different from Ethereum addresses, but Ethereum addresses and Polygon addresses (a scaling solution for Ethereum) share address spaces. The bridge may wish to publish a message so that anyone on a destination chain can be issued debt when assets are deposited into the custodian. The message may be a cryptographic puzzle, so that anyone who solves it can claim it, or only be decoded by the user who deposited the asset (e.g., knowing a hash preimage; see [14] for an example protocol using this). Advanced communicators may also have built-in mixers (see e.g., [15]), to anonymize assets as they are transferred between chains, or other unique features.

Finally, since bridges rely on communicators to relay messages, it is imperative that these entities can always write on the blockchains they communicate with. In particular, these entities should be guaranteed to be live so that a communication between chains will always occur eventually.

Communicators may be required to collect a fee in order to guarantee this liveness. First, bridges may be required to pay for transactions to submit transactions on either a source or destination blockchain (or both). For example, Ethereum transactions require gas fees, paid in Ether, which are awarded to the block producers for the chain. This provides an incentive for the inclusion of a transaction in a block and offsets the costs of block production. On other blockchains, another asset may be used as a gas fee, like an ERC-20 token. In testing or centralized solutions, gas fees may be offset by other sources (e.g., the operator of the blockchain). Second, this fee may offset the cost of operating the
communicator software. Running a node that is required to interact with a blockchain may be non-trivial. For example, to interface with Ethereum, an Remote Procedure Call (RPC) endpoint is required which may be paid service or a full Ethereum node. The former may cost per read or write to the blockchain, while the latter may be expensive to keep online and up-to-date.

III. CUSTODIAN ATTACKS

In this section, we review three exploits that have exploited the custodian component of bridges. The first exploit involves changing the privileged address that can access the digital assets, using cross-chain function calls. The second exploit aims to forge proofs that are accepted by custodians to release assets. The third exploit aims to trick the custodian into emitting deposits when it should not.

A. Truncated Function Signature Hash Collisions and Forced Transaction Inclusions

Depending on the structure of the bridge’s custodian, privileged addresses may have access to the assets in custody. This is common for centralized bridges, and is a requirement when only transactions from a particular whitelisted account are allowed to unlock these assets. This requirement is the simplest way to ensure that only an appropriate communicator can unlock funds.

Moreover, if a bridge is built in a modular way, a vault holding the assets may be separate from the contracts that are written to by the communicator directly. It may also be desirable that such a vault has its own privileged administrator.

Some bridges have the ability to execute cross-chain function calls. That is, the bridge can accept transactions on a source blockchain that include encodings of function calls to be executed on the destination blockchain (or vice versa). For example, a user may wish to bridge an asset onto another chain and immediately deposit it into a DeFi application on the destination chain. This involves calling a deposit function on the destination blockchain, if the asset is originally on the source blockchain, or vice-versa.

The goal of this exploit is to change a bridge vault’s privileged addresses to an attacker’s address, using cross-chain function calls.

The situation for this exploit is illustrated in Figure 5. In this situation, the custodian has an additional field which has special roles, in addition to receiving instructions from the communicator (possibly via another smart contract).

With this kind of bridge structure, such an exploit occurred through the following steps, illustrated in Figures 6 and 7:

1) A bridge is deployed so that anyone can call its cross-chain communication contract, specifying a function to execute. The cross-chain function call is specified via a truncated hash of the function signature; this is common for Ethereum [16], [17]. Specifically, a function signature is the first four bytes of the Keccak256 [18] hash of functionName(bytes,bytes,uint256) → 7dab77d8

Fig. 5. The structure of a Custodian with additional privileged addresses to manage the custody of assets. In this case, the cross-chain communication contract is not the contract actually holding onto digital assets, there is a Custodian contract. The custodian contract has a function changeCustodyAddress which is called by the cross-chain communication bridge bridge, and only privileged actors can move funds in the custodian contract.

functionName(bytes,bytes,uint256) → 7dab77d8

Fig. 6. An example of a cross-chain function call (top) and an example of the truncated hash of a function (bottom). Anyone can call the cross-chain communication contract, specifying a function name and its ordered argument types. An example is shown under the figure, where 7dab77d8 is the first four bytes of the Keccak256 [18] hash of functionName(bytes,bytes,uint256). This hash is used to execute a call from the other end of the bridge.

2) The attacker then defines a function such that, (a) the function has the same argument types, and (b) when the name along with the arguments taken by the function, the truncated hash is the same as calling a function changeCustodyAddress expected by the custodian. The attacker specifies a contract that they own with the function signature defined above. Finally, the attacker specifies this function and its new contract for the cross-chain execution call, resulting in this transaction’s inclusion on the destination blockchain (Figure 7).

3) The attacker notes that this transaction is now included in the destination blockchain (even if it fails), and as such can now be communicated back to the custodian on the source blockchain with proof that executing this transaction happened. The transaction is therefore able to be replayed on the source blockchain, where it succeeds and the attacker becomes a privileged actor of the vault.

The sources of the error here are steps (2) and (3).
Indeed, step (2) should be nearly impossible, as hash functions are typically assumed to be collision resistant. A hash function is collision resistant if finding two inputs to the function that result in the same output is computationally difficult to find [19]. However, as the next subsection will illustrate, implementation choices made the attack feasible in one situation. Step (3) is also a problem as a transaction’s inclusion on the destination blockchain should be insufficient to replay it on the source blockchain.

1) Real World Example: This issue was identified in the PolyNetwork bridge [20], [21]. Critically, the hash collision only required the first 4 bytes of the hashes to match. This is because the function selector, which decodes the hash, only inspects the first 4 bytes when choosing which function to call. Thus only a partial hash collision was necessary. However, finding the hash collision was necessary, but not sufficient to accomplish this attack.

In practice, the attack required running a modified communicator component, the PolyNetwork Relayer. First, the attacker sent a transaction to the destination blockchain attempting to call the function on the destination blockchain that is a hash collision corresponding to the legitimate transaction to change the custodian’s vault owner. This was communicated to the destination blockchain, included in the state tree, but ultimately did not execute correctly. This was because the custodian was not on the destination blockchain, after all; it was on the source blockchain.

However, the transaction was included on the destination chain, with proof. That is, a transaction signed by the PolyNetwork chain operator to update the vault owner was available and in the state database for the destination blockchain. The attacker was then able to force this transaction to be executed on the source blockchain, so that it was interacting with the custodian. The proof was verified (since the transaction was included on the destination chain) and the transaction executed successfully (since it was now being called on the chain on which it was intended to be called). The result was that the attacker obtained privileged roles with the vault.

2) Solution: The mitigation for this attack is to counter step (3). This is because step (2) is the well-established method for resolving function signatures on Ethereum, and likely cannot be changed without introducing compatibility issues to a bridge or a hard fork of Ethereum.

A custodian should validate that the transaction, even if it originates on the destination blockchain, provided to a custodian is legitimate. One way to do this is to ensure that the communicator cannot be bypassed. This would prevent the inclusion proof from being considered legitimate, as the attacker would not have been able to submit it to the custodian with the communicator’s signature or from the communicator’s address.

B. Incorrect Proof-of-Burn Verification

Depending on the structure of a bridge’s custodian, proofs are to be presented to the custodian in order to release assets. This type of mechanism may be common for decentralized bridges, allowing anyone with a valid proof to interact with the custodian directly for withdrawals of assets, removing the need for a centralized communicator.

The goal of this exploit is to craft fraudulent proofs that would be valid for the verification process, thereby enabling seemingly correct withdrawals.

This kind of exploit could have occurred through the following steps, illustrated in Figure 8:

1) An actor deposits funds into a custodian smart contract on the source blockchain.

2) The communicator relays this information to the debt issuer on the destination blockchain for the bridge, and the debt issuer provides the actor with a debt token.

3) The actor burns the debt token by depositing it back into the debt issuer.

4) The actor receives a so-called proof-of-burn for the token. The proof-of-burn is a string generated by the debt issuer showing that the debt token was burned.

5) The actor submits a (modified) proof-of-burn to the custodian, to unlock assets on the source blockchain, and the custodian considers the proof valid.

The source of the error here is step (5), which enables an attacker to submit a modified proof-of-burn (alongside the original proof-of-burn) to withdraw funds. The real-world exploit occurred because the proof had a leading byte that was not verified by the custodian when releasing funds, which we now describe.

1) Real World Example: This exploit was detected and patched on the Polygon/Matic Ethereum-Plasma bridge before any harm could be done [22]. We outline
the specifics of how this particular issue manifested itself for completeness, though this type of exploit may have different manifestations depending on the (incorrect) implementation of proof generation and verification for a particular bridge.

In this case, the custodian is to release funds if a proof-of-burn for the debt token is specified in a particular Merkle Patricia trie (see e.g., [23]) representing the state of the destination blockchain. In this case, the proof-of-burn includes a path to the leaf in the Merkle Patricia trie which specified that the debt token was burned (the transaction should be included when the actor submits it on the destination chain). This proof-of-burn included a branchMask parameter that should be unique.

The branchMask parameter is encoded with so-called hex-prefix encoding [2]. But at some points within the system implementation, the parameter is encoded and decoded into 256-bit unsigned integers, and during this process some information is lost. In particular, a path in the Merkle Patricia trie may have multiple valid encodings within the system.

The system was implemented to determine the path (in a trie) length encoded by a hex-prefix encoding. To use the encoding’s length, it is important to know if the length of the path is even or odd; this affects how the encoding is later expanded. The system was implemented to check that the parameter’s first nibble (4 bits) represented 1 or 3; if so, it considered the path length to be an odd number. However, it was also implemented such that in the event that the first nibble is not 1 or 3, the first byte (8 bits) is discarded but verification proceeds. Thus, there are \(2^8 - 2(2^4) = 224\) possible ways to encode a path in the Merkle Patricia trie in the situation where the first byte is discarded. In particular, there are \(2^4\) encodings for every possible bit setting of the first byte, minus the cases where the first nibble is either 1 (2\(^4\) cases; every configuration of the last 4 bits) or 3 (also 2\(^4\) cases for each configuration of the last 4 bits).

Thus the attacker would simply find a valid proof where the initial nibble was not 1 or 3, use it, and then replay the transaction for each of the remaining 223 combinations of bits for the first byte. In each case, the proof-of-burn would look legitimate and the exit would succeed, subject to delays in confirmations, delay periods, or other specific requirements of this bridge and the blockchains it connected.

2) Solution: The remedy for this exploit is correct implementation of proof verification. The original reporter of the issue notes that the first byte should in fact always be zero, reducing the number of times a valid proof-of-burn can be used to only once [22].

If the relevant proofs are built and verified correctly, this exploit will not be common, subject to common cryptography assumptions like the collision resistance of hash functions and the inability to forge digital signatures.

C. Inconsistent Deposit Logic

Bridges are often built for custom blockchains. For example, anyone developing a rollup may have a token that is used for governance or to be used as payment for gas on the rollup (instead of ETH). As a result, sometimes bridges have custom functionality for some tokens.

Moreover, a token can be “wrapped” within another token. Most commonly, Ether (ETH) is often wrapped into wrapped Ether (wETH). This is helpful because some decentralized applications do not wish to treat Ether differently from ERC-20 tokens, and wrapped Ether is an ERC-20 token. This can be helpful, as native ETH lacks a transferFrom function, among other helpful functions that are available to ERC-20 tokens. As a result, there is a wrapped Ether smart contract on the source blockchain that essentially lets anyone lock one ETH to mint one wETH.

The goal of this exploit is to trick the custodian into emitting events for deposits which are not real.

This kind of exploit occurred through the following steps, illustrated in Figure 9. It is fairly restricted in scope and requires special tokens, like wrapped Ether, to be handled differently than unwrapped assets.

1) The bridge is established in such a way that its final logic for emitting deposit events is after processing of wrapped assets. The bridge is also (incorrectly) built so that unwrapped assets allow this logic to be called, without actually supporting the transfer of those assets.

2) An attacker deposits assets into the custodian, without first wrapping the assets.

The second step is the source of the issue, and is exemplified in the real-world manifestation we now describe.

1) Real World Example: This error occurred for the Meter bridge [24]. In this occurrence, the bridge expected all assets to be transferred in a wrapped form,
Fig. 9. Two separate paths to deposit into a custodian contract.

and assumed a deposit of unwrapped assets is a mistake. However, the deposit of unwrapped assets was encoded within the same event logic that accepts wrapped assets, even though unwrapped assets were not accepted by the custodian. As a result, the custodian still emitted an event saying that funds had been transferred, even though the custodian never received them. That is, the caller continued to own their assets, but the custodian still emitted an event.

2) Solution: This particular attack is not conceptually involved. Its mere existence was enabled by a bug in the code and the branching logic. It serves as a reminder to the bridge developers that Ethereum’s native Ether is not an ERC-20 token, and both the cases of transferring Ether and its wrapped form need to be handled properly. Good engineering practices, including implementing tests, should suffice to mitigate the problem in the future.

IV. DEBT ISSUER ATTACKS

In this section we review one exploit on the debt issuer component of a bridge. The exploit aims to arbitrarily mint debt tokens on the destination blockchain.

A. Bypassing Signature Verification

The exploit aims to arbitrarily mint debt tokens on the destination blockchain. In doing so, the attacker can trade these tokens back in, honestly, and receive the corresponding assets on the source blockchain, as long as such assets are available, or for other assets on the destination blockchain.

Recall that the debt issuer smart contracts live on destination blockchain, mint debt tokens which are representations of assets on the source blockchain, and receive minting signals from a communicator (see Figure 10).

In the most straightforward implementation of debt issuers, these components mint tokens only after receiving a signed message from a communicator. This prevents unwanted tokens from being minted on the destination blockchain. To check the validity of such a signature, verification logic may be placed in a smart contract which is external to the contract issuing the debt tokens on the destination chain (see Figure 11 for an example). Moreover, if there are several communicators in a bridge, each might have its own verification logic, and modularising this logic may make sense from an engineering standpoint. This would enable verification of signatures from multiple sources, each with its own verification scheme. When a message is received in this situation, it could therefore include the address of the verification logic to be used. The logic for determining which verification contract should be used must be matched to the message, and including the address of a contract that implements the verification logic is a straightforward implementation. However, problems arise if matching allows messages to be matched with arbitrary verification logic, as we now describe.

This exploit was executed using the following steps, illustrated in Figure 12:

1) An attacker deploys a smart contract on the destination blockchain that has a function that the debt issuer expects to call to verify a signature. The function is implemented so that any signature is “verified”, possibly by implementing the verification function so that it always returns true. This verification contract therefore accepts any string as a valid signature.

2) The attacker from step (1) sends a debt issuance signal to the debt issuer, referencing the smart contract they deployed in step (1) as the verification logic.

3) The debt issuer provides debt tokens to the attacker.

The source of the issue here is in steps (2) and (3).
The debt issuer should not have accepted just any smart contract as verification logic. After the attacker gains the debt tokens, they can behave honestly to bridge the assets back to the source blockchain, stealing funds from the custodian.

1) **Real World Example:** Unfortunately, this situation was exploited on the Wormhole bridge [25]. The result was that about 120,000 Ether was minted on Solana, which was worth about $323 million USD at the time the exploit occurred. Much of this Ether was transferred back to Ethereum.

2) **Solution:** The example in Section IV-A1 was enabled by the attacker’s ability to provide both the signature (used to confirm the authenticity of the transaction) and the reference to the signature verifier (used to confirm the signer’s authorization to issue the transaction) within the user transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

V. COMMUNICATOR ATTACKS

In this section we review two exploits targeting the communicator component of a bridge. The first exploit aims to trick the communicator into forwarding invalid messages from one blockchain to the next, while the second uses a 51% attack on a blockchain to cause a blockchain re-organization after the communicator receives a valid message. These exploits can be thought of as polluting the data source of an oracle, the communicator.

A. **Forwarding Invalid Messages**

The goal of this exploit is to trick the communicator into forwarding invalid messages from the source blockchain. This will result in incorrect debt issuance on the destination blockchain, minting debt tokens that are not mapped to assets in custody on the source blockchain.

The exploit proceeded according to the following steps, illustrated in Figure 13:

1) The bridge is established in such a way that its communicator watches events emitted from the source blockchain. The communicator watches for these events on transactions that deal with a particular address, namely, the address of the custodian for the bridge. Notably, it watches all events in such a transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

2) **Solution:** The example in Section IV-A1 was enabled by the attacker’s ability to provide both the signature (used to confirm the authenticity of the transaction) and the reference to the signature verifier (used to confirm the signer’s authorization to issue the transaction) within the user transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

V. COMMUNICATOR ATTACKS

In this section we review two exploits targeting the communicator component of a bridge. The first exploit aims to trick the communicator into forwarding invalid messages from one blockchain to the next, while the second uses a 51% attack on a blockchain to cause a blockchain re-organization after the communicator receives a valid message. These exploits can be thought of as polluting the data source of an oracle, the communicator.

A. **Forwarding Invalid Messages**

The goal of this exploit is to trick the communicator into forwarding invalid messages from the source blockchain. This will result in incorrect debt issuance on the destination blockchain, minting debt tokens that are not mapped to assets in custody on the source blockchain.

The exploit proceeded according to the following steps, illustrated in Figure 13:

1) The bridge is established in such a way that its communicator watches events emitted from the source blockchain. The communicator watches for these events on transactions that deal with a particular address, namely, the address of the custodian for the bridge. Notably, it watches all events in such a transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

2) **Solution:** The example in Section IV-A1 was enabled by the attacker’s ability to provide both the signature (used to confirm the authenticity of the transaction) and the reference to the signature verifier (used to confirm the signer’s authorization to issue the transaction) within the user transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

V. COMMUNICATOR ATTACKS

In this section we review two exploits targeting the communicator component of a bridge. The first exploit aims to trick the communicator into forwarding invalid messages from one blockchain to the next, while the second uses a 51% attack on a blockchain to cause a blockchain re-organization after the communicator receives a valid message. These exploits can be thought of as polluting the data source of an oracle, the communicator.

A. **Forwarding Invalid Messages**

The goal of this exploit is to trick the communicator into forwarding invalid messages from the source blockchain. This will result in incorrect debt issuance on the destination blockchain, minting debt tokens that are not mapped to assets in custody on the source blockchain.

The exploit proceeded according to the following steps, illustrated in Figure 13:

1) The bridge is established in such a way that its communicator watches events emitted from the source blockchain. The communicator watches for these events on transactions that deal with a particular address, namely, the address of the custodian for the bridge. Notably, it watches all events in such a transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

2) **Solution:** The example in Section IV-A1 was enabled by the attacker’s ability to provide both the signature (used to confirm the authenticity of the transaction) and the reference to the signature verifier (used to confirm the signer’s authorization to issue the transaction) within the user transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

V. COMMUNICATOR ATTACKS

In this section we review two exploits targeting the communicator component of a bridge. The first exploit aims to trick the communicator into forwarding invalid messages from one blockchain to the next, while the second uses a 51% attack on a blockchain to cause a blockchain re-organization after the communicator receives a valid message. These exploits can be thought of as polluting the data source of an oracle, the communicator.

A. **Forwarding Invalid Messages**

The goal of this exploit is to trick the communicator into forwarding invalid messages from the source blockchain. This will result in incorrect debt issuance on the destination blockchain, minting debt tokens that are not mapped to assets in custody on the source blockchain.

The exploit proceeded according to the following steps, illustrated in Figure 13:

1) The bridge is established in such a way that its communicator watches events emitted from the source blockchain. The communicator watches for these events on transactions that deal with a particular address, namely, the address of the custodian for the bridge. Notably, it watches all events in such a transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

2) **Solution:** The example in Section IV-A1 was enabled by the attacker’s ability to provide both the signature (used to confirm the authenticity of the transaction) and the reference to the signature verifier (used to confirm the signer’s authorization to issue the transaction) within the user transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

V. COMMUNICATOR ATTACKS

In this section we review two exploits targeting the communicator component of a bridge. The first exploit aims to trick the communicator into forwarding invalid messages from one blockchain to the next, while the second uses a 51% attack on a blockchain to cause a blockchain re-organization after the communicator receives a valid message. These exploits can be thought of as polluting the data source of an oracle, the communicator.

A. **Forwarding Invalid Messages**

The goal of this exploit is to trick the communicator into forwarding invalid messages from the source blockchain. This will result in incorrect debt issuance on the destination blockchain, minting debt tokens that are not mapped to assets in custody on the source blockchain.

The exploit proceeded according to the following steps, illustrated in Figure 13:

1) The bridge is established in such a way that its communicator watches events emitted from the source blockchain. The communicator watches for these events on transactions that deal with a particular address, namely, the address of the custodian for the bridge. Notably, it watches all events in such a transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.

2) **Solution:** The example in Section IV-A1 was enabled by the attacker’s ability to provide both the signature (used to confirm the authenticity of the transaction) and the reference to the signature verifier (used to confirm the signer’s authorization to issue the transaction) within the user transaction. As a result, the attacker was able to authorize any calls via a friendly custom verifier. Therefore, a clear prevention of the attack is ensuring that verifiers cannot be provided by users. Verifiers need to be absolutely trusted elements of the system, and as such, can be deployed only by trusted entities, and users cannot be provided with an option to choose a dishonest verifier to authorize their transaction.
The inherent cause of the attack is the computational power in the network and the duration of the attack. More computational power and a longer attack duration increase the cost, while shorter attacks are cheaper. For blockchains which use different methods to append blocks, like those which use so-called proof-of-stake (see e.g., [28]) consensus algorithms, the cost of this attack will depend on different factors. For example, proof-of-stake systems may have an increased risk for this attack if there are too few validators who propose blocks, the validators can be bribed or collude, or if the cost required by stakers is too low.

The exploit would proceed according to the following steps:

1) An attacker honestly deposits funds into a bridge via the source blockchain’s custodian, which issues debt on the destination blockchain.

2) After waiting a small-but-not-too small amount of time (enough for several confirmations of the transaction, perhaps about 5; this is about 15 additional blocks on Ethereum at the time of writing), the attacker rents computational power to enact a moderately long 51% attack (about 1 hour). This attack establishes another chain which is canonical after the attack but does not have the attackers transaction from step (1) included on the chain.

The source of the exploit here is contained in step (2), where the attacker re-organizes the source blockchain but also has the issued funds on the destination blockchain. This attack is a specific instance of double-spending a token [8].

1) Real World Example: This exploit has not yet been executed. Prior to Ethereum’s switch to proof-of-stake consensus (August 2022), the website Crypto51.app [30] reported that the cost of a 1-hour long 51% attack on Ethereum would cost $600,374 USD. This value is lower for chains that have less computational power associated with them. Nevertheless, as the real-world examples of bridge attacks referenced in this paper demonstrate, the attacker’s profit often exceeds this number.

2) Solution: The inherent cause of the attack is the bridge’s wrong assumptions about the finality of blocks. The bridge needs to ensure that if a reorganization of the source chain happens and the deposit transaction becomes invalidated, the same invalidation happens on the target chain. This is a difficult task for bridges that are not implemented and operated natively by the target blockchain itself, and reside on it in the form of a third party application. The native bridges may implement a mechanism that keeps track of deposit nonces and the total bridged value on the source chain, and subsequently require “committing” the nonce and value sequence in the transactions that release the assets from the custodian. The nonces would have to have a fixed sequence that prohibits skipping (e.g., integers that increment by 1 with every deposit) so that they guarantee that if a
deposit transaction is dropped, or it value changes, all the subsequent deposits to the target chain and withdrawals from it become invalid (i.e., result in failed transactions). Consequently, the bridge would need to ensure that such a reorganization is properly reflected on the target chain so that the users whose assets were now not deposited to the target chain or withdrawn have their balances adjusted accordingly, and the integrity of the asset amounts between the two chains is not violated.

It is also important to note that the roles of a source and target chain are to a certain extent symmetric—a chain that is in the role of a source during the deposit may be in the role of a target during the withdrawal. While a bridge may be a native component of one of the chains and may ensure that the chain can respond to the reorganization of another chain, it is unlikely that it would be able to guarantee such a reorganization for both the chains, in particular, for the Ethereum mainnet.

The prevention of the attack described in this section is a difficult problem and it would make for a great subject for future research.

VI. TOKEN INTERFACE ATTACKS

In this section we describe some exploits based on the token interfaces used in bridges. The first exploit relates to token approvals for bridges, while the second exploits the EIP-2612 [31] interface function built into some ERC-20 tokens.

A. Infinite Approvals and Arbitrary Executions

The goal of this exploit is to take user’s funds directly, rather than stealing them from the custodian or debt issuer, by leveraging bridge components which can call other smart contracts.

A valuable use case of ERC-20 tokens is the ability to approve others to spend your tokens. For example, you may wish to approve a bridge to spend your tokens, so that in the future if you use a decentralized application to interact with the bridge, it can take funds on your behalf, through the decentralized application you are interacting with. This is achieved by having the user call approve (possibly specifying some specific amount) on the token, listing the bridge’s relevant smart contract address, and later having the bridge call transferFrom to take funds from the user. The latter call will only succeed if the user has approved the bridge to act on the user’s behalf.

Due to large gas concerns, users often grant applications and bridges infinite approvals. This is because the approve call is a transaction that must be executed on chain, for which gas must be paid. As a result, an infinite approval removes the requirement of subsequent approval calls, saving the user gas fees in the future.

Moreover, recall from Section III-A, due to the composability of smart contracts (especially popular in DeFi applications), bridges often call other smart contracts directly. To do this, a bridge may have an arbitrary function execute which takes an ABI encoded description of the function to call (see also Section III-A and [17]).

This exploit proceeded according to the following steps, illustrated in Figure 14:

1) A user provides a bridge that can call smart contracts with an infinite approval to a token.
2) An attacker calls execute with an encoding of transferFrom to take the honest user’s tokens, rather than their own. This succeeds since the bridge executes the transferFrom call, and it has approval to take the user’s tokens. However, since the attacker initiated the call, the debt issuer on the destination chain issues the debt in the name of the attacker.

The source of the error for this exploit is in step (2), in which the debt is incorrectly issued to the wrong party. After the attacker receives the debt token on the destination blockchain, they can bridge the asset back to the source blockchain and withdraw the funds. The user is now powerless to recover those tokens.

1) Real World Example: This exploit was possible on an earlier version of the Multichain (formerly “AnySwap”) project [32]. The attack vector was documented before it was exploited, and the finders were awarded a $1,000,000 USD bounty for finding and reporting the issue, which they first demonstrated on a local fork of Ethereum. At the time the exploit was reported, almost 5,000 accounts had granted infinite approval to the bridge in question.

2) Solution: As the vulnerability is enabled by the bridge issuing debt tokens to a user who did not supply the tokens on the source chain, one possible remedy is to ensure that the debt is always issued to the account that provided the tokens on the source chain. However, this strongly limits the design of the bridge and may cause problems when bridging tokens in the custody of
a smart contract. A specific concern with this solution is the use of a so-called multisig wallet (see e.g., [33]) — a smart contract that holds tokens on behalf of multiple users whose joint signatures are required for releasing such tokens. Such a smart contract may be available on the source chain, however, due to how the smart contract addresses are determined (see [2]), such a multisig wallet may not be available on the target chain. Other measures, such as disallowing generic calls to functions such as execute, may negatively impact the required features of the bridge, and thus do not appear viable without disrupting the business logic.

B. Permits and Non-Reverting Fallback Functions

Similar to Section VI-A, the goal of this exploit is to take user’s tokens directly, rather than stealing them from the custodian or debt issuer, by leveraging bridge components which can call functions in poorly implemented token contracts.

Some ERC-20 tokens have a permit function, which enables a user to sign a message enabling others to use one’s tokens; these implement EIP-2612 [31]. A message is signed using the account’s private key. These messages are not transactions (they are not executed on-chain), and do not use gas; as a result, they are attractive in some settings as they are free for users. Once someone has obtained a signed permit message, they can go to the smart contract for the token and call a function to verify the signed message. If the verification succeeds — that is, a verify function for the permit does not revert — the holder of the permit is approved to use the signer’s tokens (e.g., via the transferFrom function). A function reverts on Ethereum if it runs out of gas, an error occurs, or an assertion (written as a require or assert) fails in the code being executed. The permit holder can obtain the approval by calling a function redeemPermit and “using up” the permit. The fact that the verify permit function is expected to revert if the verification fails is key, as we explain next.

The expectation that a function reverts is problematic when developers are not aware of the expectation. Smart contracts on Ethereum have a fallback function, which is called whenever a function that is supposed to be called on a smart contract cannot be found (i.e., it is not implemented). If an ERC-20 smart contract implements a fallback function that does not revert, every time a function that does not exist within the implementation is called, the call will succeed. This can be problematic, as we now exemplify in this exploit.

The exploit proceeded according to the following steps, illustrated in Figure 15:

1) A user wishes to bridge tokens to another chain from Ethereum where (a) the bridge supports permit redemption for tokens, and (b) the bridge has custody of at least one token that does not implement the EIP-2612 permit functions but implements a fallback function that does not revert.

2) The user gives the bridge infinite approvals (and may or may not successfully send some tokens over the bridge).

3) An attacker asks the bridge to redeem a string, claiming it is a permit, for approval on the token used by the honest user in steps (1) and (2). The bridge attempts to verify the supplied string; as the verify function is not implemented for any permit but the fallback function never reverts, the bridge accepts the permit. Next, the bridge contract calls the redeem function for the permit, which is not implemented either, but since the fallback function never reverts, the bridge thinks it has succeeded in obtaining the approval. The bridge calls transferFrom on behalf of the attacker onto the bridge, and issues debt in the attacker’s name.

As in the Section VI-A, the source of the error for this exploit is in step (3), in which the debt is incorrectly issued to the wrong party. After the attacker receives the debt token on the destination blockchain, they can bridge the asset back to the source blockchain and withdraw the funds. However, this reasoning is different: the bridge was not implemented to check that verify was actually called, rather than the fallback function.

1) Real World Example: The situation above was described as a feasible bug in the deprecated Polygon Bridge Zap [34]. The operators fixed the issue in subse-
quent versions even before it was discovered, but because a version was already deployed on the blockchain the exploit happened. To mitigate this, the bridge operators submitted transactions to withdraw funds themselves, but were front-run by an arbitrage bot. However, the arbitrage bot later returned the front-run profit when the bot operator learned that the transaction was executed with good intention. The vast majority of funds were held in escrow by the Polygon team after this effort, resulting in no significant loss of funds from the bridge.

2) Solution: The core idea of the vulnerability lies in the fact that the bridge calls a non-existent function on a token to redeem the permit, and that the token’s implementation allows calls to non-existent function without reverting. In order to eliminate it, the bridge needs to break this condition. This means that the bridge needs to be aware of the implementation of the token, and avoid attempts to use the permit logic if the token does not implement it, or does not implement it correctly. As the Ethereum blockchain does not allow for checking interface and implementation on-chain, a list of tokens maintained by the operator of the bridge or an off-chain logic would have to be used for detecting whether the permit mechanism should be available.

VII. RELATED AND FUTURE WORK

To our knowledge, there is no systematization of the attacks that have occurred on bridges in recent years. Others have studied these components, though they largely do so from a theoretical point of view, rather than reviewing the concrete faults of previous systems. We first list a few of the relevant works.

McCorry et al. [35] review the literature involving bridges and provide a detailed breakdown of roles and components. Although their terminology differs from ours, the key concepts they identify for bridges are compatible with the major components we define in Section II. Their work provides a more fine-grained overview; for example, they dive into concepts like the operator of a communicator (i.e., whether it is centralized, involves a multi-signature wallet, or purely trustless), protocol assumptions (e.g., expanding properties beyond liveness), and things like rate-limiting transactions for the bridge. They offer research directions for improvements based on various assumptions and dilemmas that bridges aim to solve or sidestep. However, their threat models are high level and theoretical; they are not reviews of attacks that have occurred in practice. For example, they emphasize the need to prevent censorship of communicators. This is critical for bridge design and complementary to the review of concrete issues we review in this work. This is discussed somewhat less formally for layer two solutions by L2Beat.com [36].

There has been much work on cross-chain communication itself. Zamyatin et al. [37] study communication between blockchains, necessary for the communicator component of a bridge. They look at which assumptions are required, classify and evaluate existing cross-chain communication protocols, and generalise the systems we describe in this work. They list challenges that must be overcome for safe and effective cross-chain communication. These challenges show that bridge communication is difficult and indicate why bridges are so complex and therefore prone to implementation errors. Specific instances of cross-chain communication that involve proving the state of one ledger to another are explored in related works like [38] and [39]. Other systems like Cosmos [40] and Polkadot [41] aim to support cross-chain communications and transfers as core functionality.

Wang [42] and Han et al. [43] both describe possible frameworks to consider the security of cross-chain systems, like bridges, but do not include detailed attack descriptions. Haugum et al. [44] list several theoretical concerns and attacks (e.g., double spending), but do not describe actual failures in these categories in any detail. These works are complementary to ours in that our work provides detailed examples that can be used to justify the properties of cross-chain systems that these authors outline.

For future work, it would be interesting to study preventative measures for these attacks. While many attacks presented are implementation specifics, we wonder if a general framework or set of standards can help mitigate these issues. In particular, custodian or debt issuer standards may reduce errors with cross-chain calls or decrease erroneous event emissions. Such a standard could be an interface akin to the ERC-20 or ERC-721 standards. There are some initial efforts to standardize the transfer of digital assets (e.g., [45]), though they do not yet appear final or adopted.

Moreover, a wishlist of specific properties for security should be explicated. The related work targets the high level properties, but is insufficient to guide new bridge developers. These high level properties like liveness may be fairly obvious, but many of the attacks reviewed in this paper may fade into obscurity and become unknown to new community members. We fear that new members may repeat these issues, and hope that a guideline for bridge construction may be established which would improve the quality of future bridges.

VIII. CONCLUSION

We explained several bridge attacks and suggested mitigations for most of them. Our work should not be seen as an exhaustive list of security issues to prevent, but rather as an insightful survey of principles that were exploited in the past that the developers of bridges should keep in mind. Our work is an illustration of the complexity of the bridge systems and the size of the attack surface that they expose for potential exploitation.
[42] Gang Wang. SoK: Exploring blockchains interoperability. *IACR Cryptol. ePrint Arch.*, page 537, 2021.

[43] Panpan Han, Zheng Yan, Wenxiu Ding, and Zhiguo Wan. A survey on cross-chain technologies. *Distrib. Ledger Technol.*, Jan 2023.

[44] Terje Haugum, Bjørnar Hoff, Mohammed Alsadi, and Jingyue Li. Security and privacy challenges in blockchain interoperability - A multivocal literature review. In Mirosław Staron, Christian Berger, Jocelyn Simmonds, and Rafael Prikladnicki, editors, *EASE 2022: The International Conference on Evaluation and Assessment in Software Engineering 2022, Gothenburg, Sweden, June 13 - 15, 2022*, pages 347–356. ACM, 2022.

[45] Secure Asset Transfer Protocol. https://datatracker.ietf.org/wg/satp/about/.