

DEFI SECURITY SUMMIT 2026 · MUMBAI · 1–2 NOVEMBER

Security Starts Before the Signature

Trustless Transaction Analysis for Ethereum Wallets

DON'T TRUST. VERIFY!

Steffen Kux · Co-CEO & Co-Founder, corpus.core – colibri.stateless – transaction security agent (TSA)

Uniswap

Nonce #47

1 **Review what you will sign**

Signing is an irreversible action so make sure you know what you are signing. Read more.

2 **Compare with your wallet**

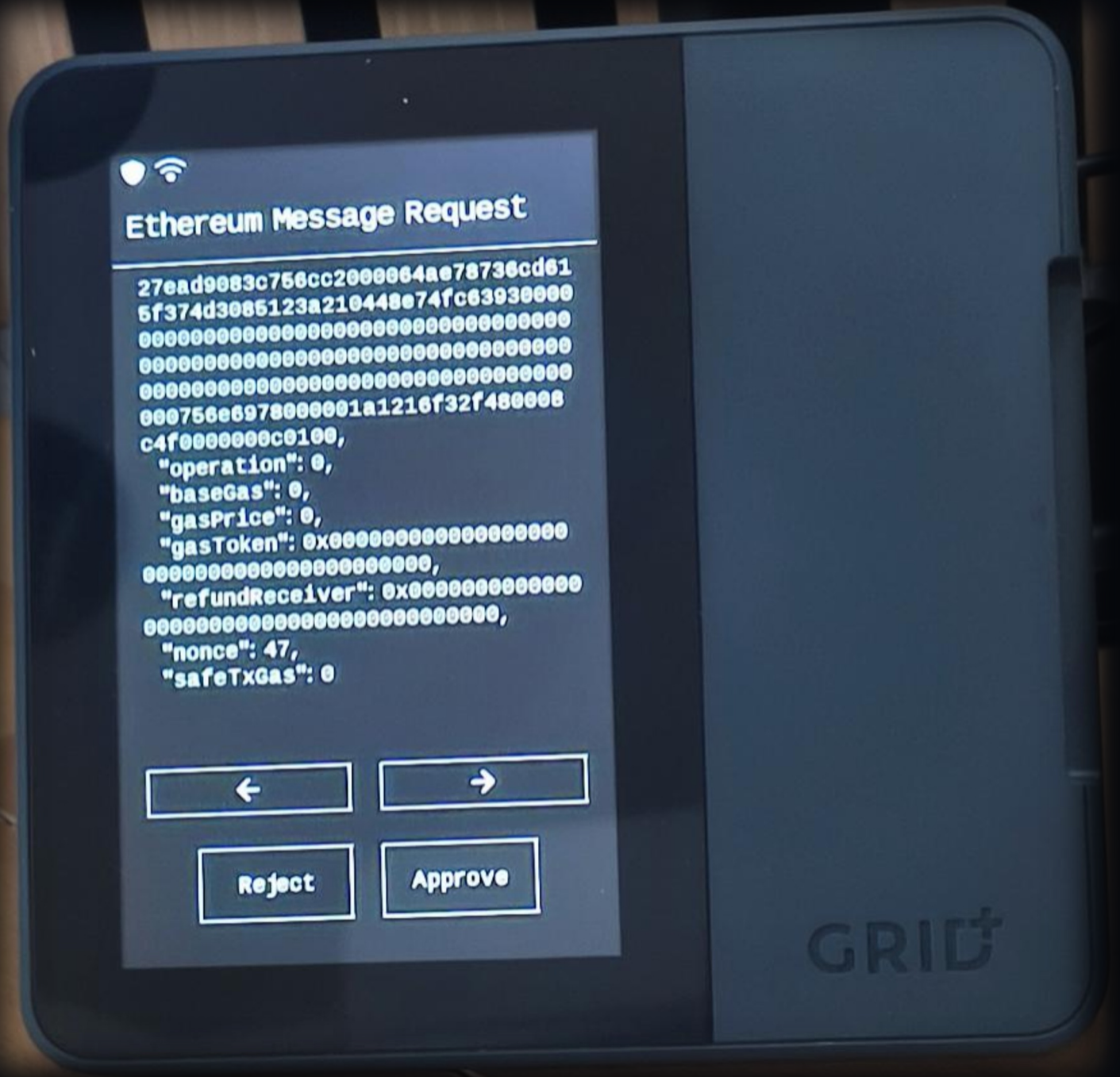
Once you click **Sign**, the transaction will appear in your signing wallet. Make sure that all the details match.

3 **Verify with external tools**

You can additionally cross-verify your transaction data in a third-party tool like Safe Utils.

Data | Hashes | JSON

To	UniversalRouter
	0x23617e59A5925b2A4Bf75d73fff
Value	1000000000000000000
Data	0x3593564c000000000000000000000000000000000000000000000000000000000000006000000000000000000000000000000000000000000000000000000000


Ethereum Message Request
27ead9083c756cc2000064ae78736cd61
5f374d3065123a210448e74fc63930000
000000000000000000000000000000000
000000000000000000000000000000000
000000000000000000000000000000000
000756e697000001a1216f32f480008
c4f0000000c0100,
"operation": 0,
"baseGas": 0,
"gasPrice": 0,
"gasToken": 0x0000000000000000
0000000000000000000,
"refundReceiver": 0x0000000000000
0000000000000000000000000,
"nonce": 47,
"safeTxGas": 0
Reject
Approve
GRID+

0x3593564c0000000000000000000000000000000000000000000000000000000060000
0000000000000000000000000000000000000000000000000a000000000000000
0000000000000000000000000000000006ac9194b000000000000000000000000
0000000000000000000000000000020b000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000
0000000200000000000000000000000000000000000000000000000400000
0000000000000000000000000000000000000000000000000a000000000000000
0000000000000000000000000000000040000000000000000000000000000000
0000000000000000000000000200000000000000000000000000000000000000
0000000de0b6b3a764000000000000000000000000000000000000000000000000
000000140000000000000000000aa290a3c526d12b244d5926e2de1dfb7e28a931c0000
0000000000000000000000000000000000000de0b6b3a764000000000000000000000
0000000000000000000000000a0a3ad4f172cda3000000000000000000000000000000
0000000000000000000000000000c000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000
000001200000000000000000000000000000000000000000000000002bc02aa
a3332abfe8a710e5c4a87ead9083c756c10081064ae78736cd6a1b974d3085123a21044ac9afc
6393000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000756e6978000001a2ca8f32f480008c4f0000000c0100

01 · THE PROBLEM

The Display Is the Attack Surface

Bybit, February 2025: $1.5 billion signed away with valid signatures

WHAT THE SIGNERS SAW

A routine transfer from the cold wallet, in a familiar UI.

WHAT THEY ACTUALLY SIGNED

A delegatecall that replaced the wallet's logic.

dApp builds tx → Wallet displays → User trusts → **Sign**

Not a Bybit problem: three trust assumptions before every irreversible signature.

02 · TODAY'S FIX

Ask Someone Else

Transaction simulation today is a remote service

Wallet
Sends unsigned tx and your address

→

Simulation service
Its node, its state, its code

→

"Looks safe"
You sign on its word

NEW TRUST ASSUMPTION
The service, its node and its result must be honest and uncompromised.

NEW LEAK
Your intent and your wallet reach a third party before you sign.

Better than nothing. But still blind, with one more party to trust.

03 · THE SHIFT

Signing Is Not Enough

From signing transactions to understanding consequences, locally

TODAY

Create → Display → Sign

WHAT WE NEED

Create → Verify → Simulate → Analyze → Explain → Decide → Sign / Reject

Four new steps (blue) and a decision point (gold), all on the user’s device.

DON'T TRUST.

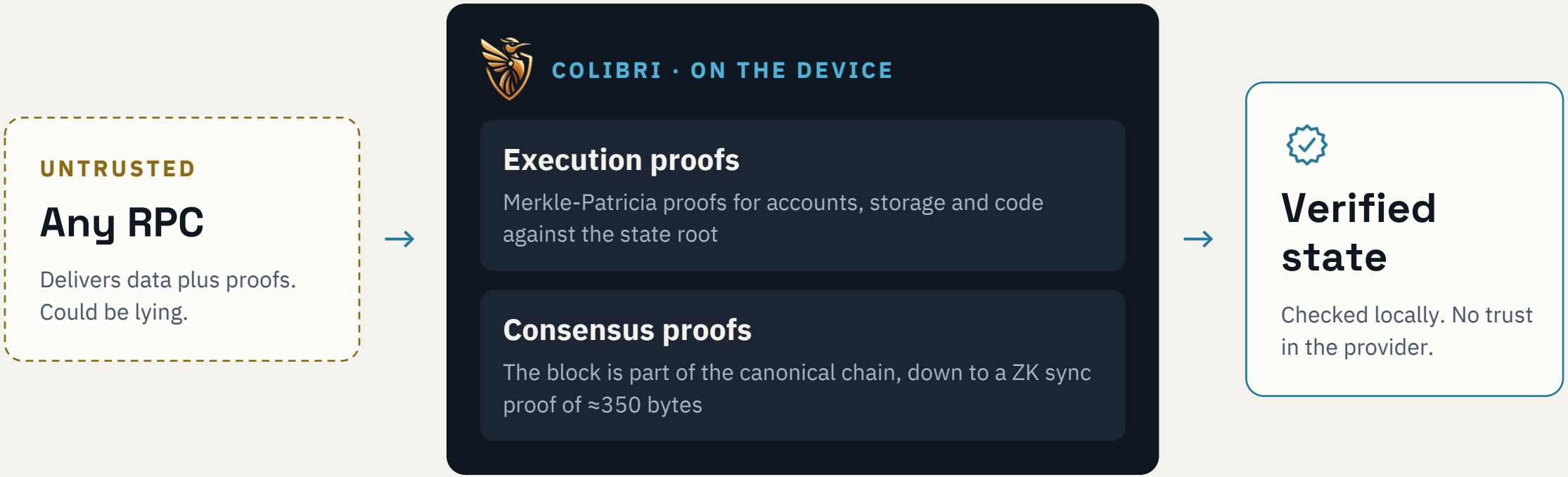
VERIFY!



04 · STEP ONE: VERIFY

Don’t Trust the RPC

If the state lies, the simulation lies. Proofs turn RPC data into verifiable inputs.



05 · STEP TWO: SIMULATE

Execute Before You Sign

Local simulation on verified state

Unsigned transaction
to · value · calldata

Verified state
every slot it touches, with a proof

→

</>

Local EVM
on the user's device

→

Calls
incl. delegatecalls

Transfers
ETH, tokens, NFTs

Approvals
who may spend

State
owner, impl. slots

We don't ask anyone what will happen. We compute it ourselves.

06 · THE NEXT PROBLEM

Humans Don’t Read Execution Traces

Knowing what happens is not the same as understanding it

```
CALL 0x3fC9…a81E execTransaction(…)
DELEGATECALL 0x9622…7d21 0xa9059cbb(…)
SSTORE slot 0x00: 0x34cf…11a2 → 0x9622…7d21
LOG ExecutionSuccess(0x5e1b…, 0)
balance Δ 0x3fC9…a81E: 0
gas used: 71 342
```

Correct.
Verified.
Unreadable.

The critical line is there, but nobody at a confirm dialog will find it.

From Consequences to Intent

Does the transaction do what I actually wanted?

USER INTENT

Swap 1 ETH → USDC

SIMULATED EFFECTS (EXAMPLE)

Effect	Matches intent?
–1 ETH from your wallet	Yes, expected
+USDC to your wallet	Yes, expected
approve(USDT, 0x7a3..., unlimited)	No, unexpected
0.05 ETH sent to 0x7a3...	No, unexpected

The question is no longer “is this valid?” but “is this what I meant?”

08 · STEP FOUR: EXPLAIN

Let AI Explain, Not Verify

Deterministic facts first. Probabilistic interpretation second.

DETERMINISTIC

What is true

Proofs and the EVM: verified state, executed effects. Reproducible.

PROBABILISTIC

What it means

AI and rules: explanation, risk signal, recommendation. Never the source of truth.

And it runs locally: intent, balances, positions and code never leave the device.

09 · DOMAIN SLMS

Small Models, Narrow Expertise

Frontier quality on one task, small enough to run locally

	Frontier LLM	Domain SLM
Explaining transactions	Strong	Strong
Everything else	Strong	Weak, by design
Size and location	Huge, in a data center	Small, in your browser or app
Your data leaves the device	Yes	No

Your wallet doesn’t need a model that writes poems and knows every celebrity breakup from the tabloids. It needs one that knows what this calldata does to your money.

10 · THREAT MODEL

Threat Model and Limits

What pre-sign analysis can and cannot guarantee

Covered

- A lying RPC or manipulated state
- A compromised frontend or display
- Hidden calls, approvals and transfers
- Unexpected state changes, e.g. upgrades

Limits

- State can change before inclusion
- A compromised device defeats any local check
- The analysis is only as good as the known intent

Not a guarantee of safety. A guarantee that you see what you sign.

11 · PUTTING IT TOGETHER

The Transaction Security Agent

Verify → Simulate → Analyze → Explain → Act

In: user intent

In: unsigned transaction

In: verified RPC data

In: relevant contract code

Verify
Colibri

Simulate
Local EVM

Analyze
Effects vs. intent

Explain
Local SLM + rules

Act
Warn, recommend

⚠ **High risk.** Besides your swap, this transaction grants unlimited USDT to an unknown address. Recommendation: **reject**.

DON'T TRUST.

VERIFY!

DON'T SIGN BLIND.

GET INVOLVED

Try Colibri

Whitepaper



whitepaper.colibri-stateless.tech

Specification



specification.colibri-stateless.tech

GitHub Repository



repository.colibri-stateless.tech

corpuscore.tech X: @CorpusCoreHQ · @ColibriClient



Thank you very much!

Steffen Kux

Co-CEO & Co-Founder, corpus.core

https://corpuscore.tech