DeFi data tracing using ML

Frédéric Dupont-Marillia

Autorité des marchés financiers

LAMBDA

Laboratoire analytique, modélisation et BI de L'AUTORITÉ DES MARCHÉS FINANCIERS

Agenda

Context

- Who are we
- DeFi tracing
- About blockchain

The problem

- Main objective
- Secondary objective

Technical details

- DeFi transactions
- Smart contracts
- Examples
- Our approach
- Finally

Who are we?

The Autorité des Marchés Financiers (AMF) **du Québec** is the regulator for the following areas:

- Insurance
- Securities
- Deposit institutions (banks and credit unions),
- Derivative instruments
- Financial products
- Exchanges (ex: stock market)
- **Cryptocurrencies**

Among its missions in the DeFi field:

- Cyber-Investigations
- Monitoring crypto markets

Autorité des marchés financiers

DeFi tracing

Cyber-Investigations

In many investigations, the Authority must analyze transactions between wallets in order to:

- Determine the amount of fraud
- Identify the stakeholders
- Highlight the mechanisms of manipulation

Data analyzed are

- Transactions
- Addresses
- Pools
- Smart contract codes

Autorité
des marchés
financiers

DeFi tracing

What are the questions we try to answer:

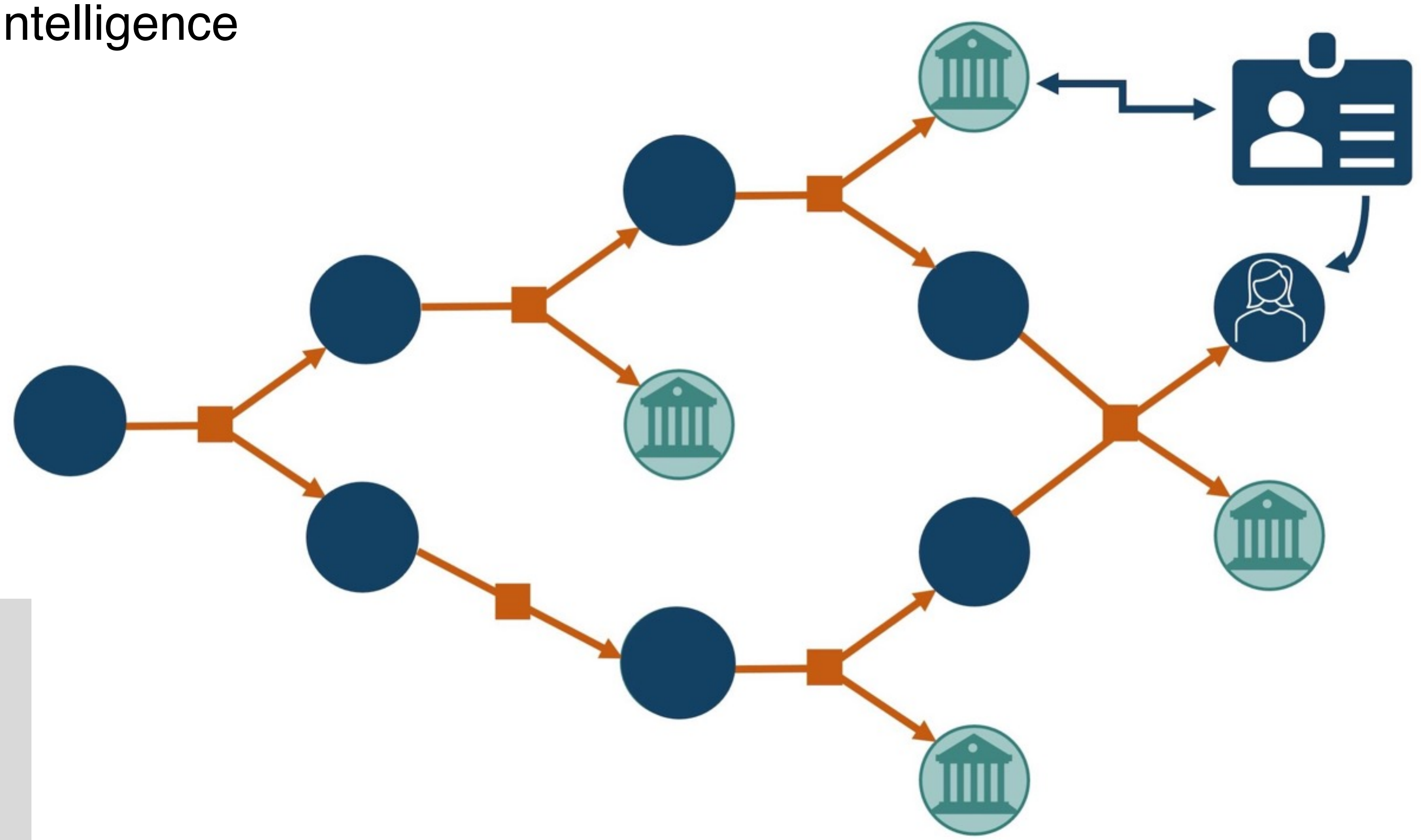
Who are the most profitable wallets?

How were they profitable?

Who own the liquidities?

It is correlated with the rest of the market?

Track assets until we can match them with other intelligence



Need

We need to analyze and process a large amount of transactions

About Blockchain

Transparency

Blockchain networks are an “**open book**”, generally providing each node with a complete copy of the network’s database.

Traceability

The chronological recording of transactions allows users to track the chain of ownership of assets recorded in the database.

Immutability

The distributed nature of blockchain databases means that information is permanently registered and resistant to tampering.

About Blockchain

Transparency

Blockchain networks are an “**open book**”, generally providing each node with a complete copy of the network’s database.

Traceability

The chronological recording of transactions allows users to track the chain of ownership of assets recorded in the database.

Immutability

The distributed nature of blockchain databases means that information is permanently registered and resistant to tampering.

Complexity

The blockchain is in permanent evolution. Its technology is getting more and more complexes to improve scalability and flexibility of its ecosystem.

Main objective

Develop a machine learning model capable of tracing transactions

Model inputs

Data extracted from the blockchain

Model outputs

Amounts, tokens exchanged, fees and values.

Key Success Factor: A key factor for success in this mission is the reliability of the tracing. A tracing reliability score could therefore represent a very important added value.

Secondary objective

Improve dataset generation to improve process robustness

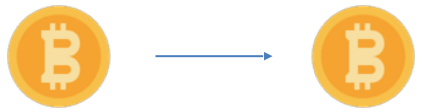
Possible avenues

- Use smart contracts
- Generate synthetic data

Key Success Factor: create a method that is scalable and can evolve with technology evolution and code updates.

DeFi transactions

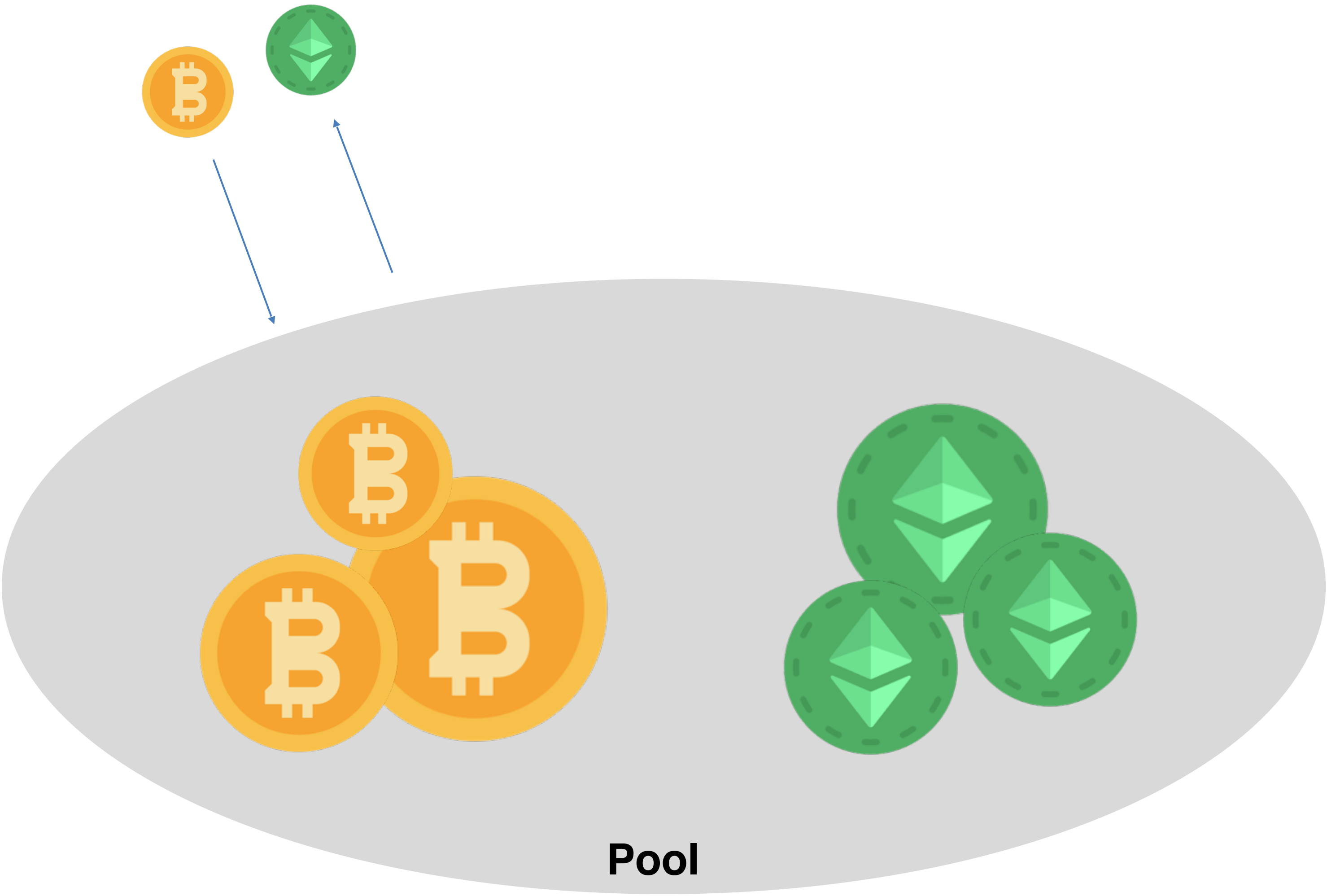
Different types of interactions

- Transfers 

DeFi transactions

Different types of interactions

- Transfers
- Swaps

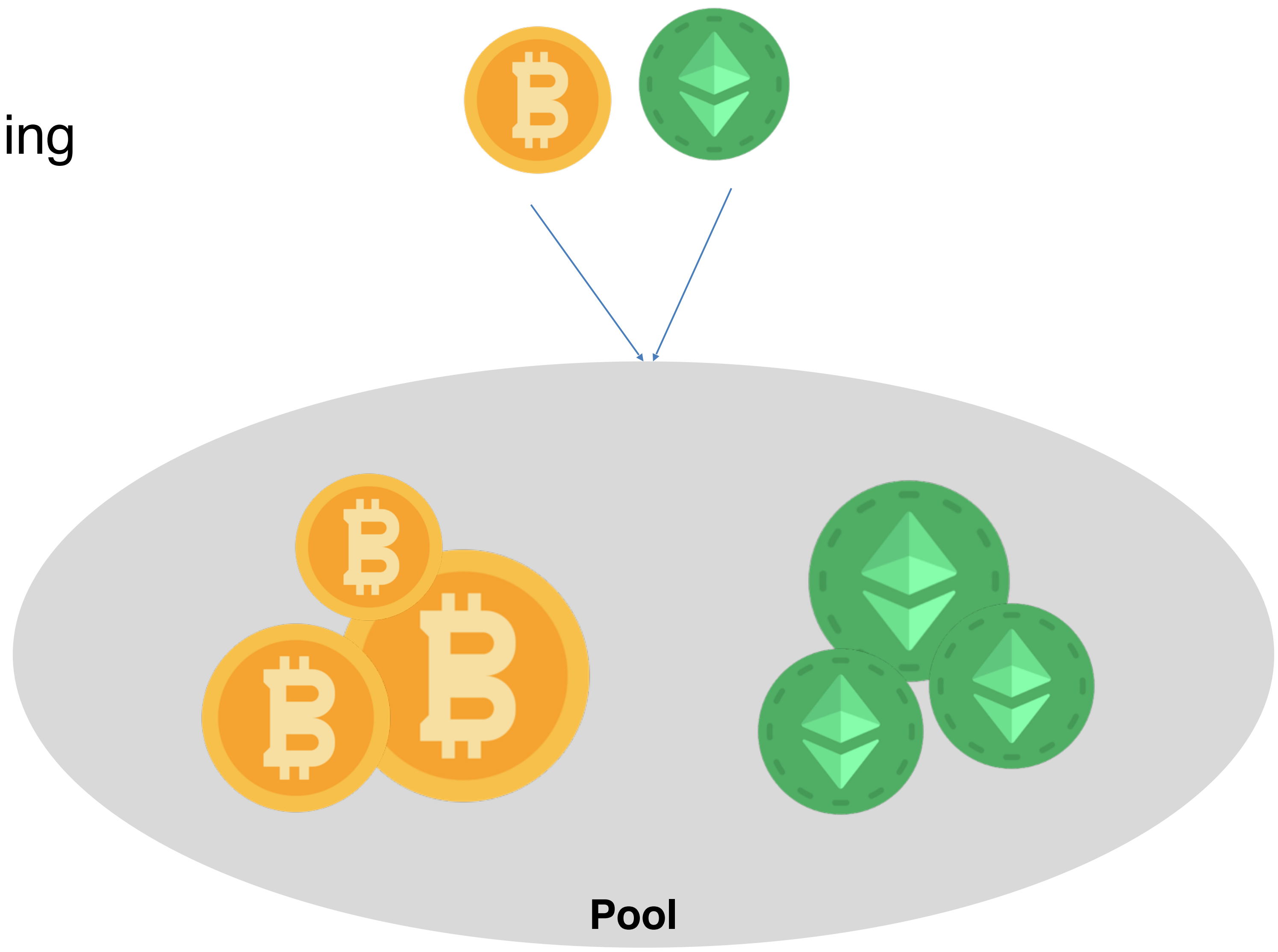


Autorité des marchés financiers

DeFi transactions

Different types of interactions

- Transfers
- Swaps
- Liquidity providing



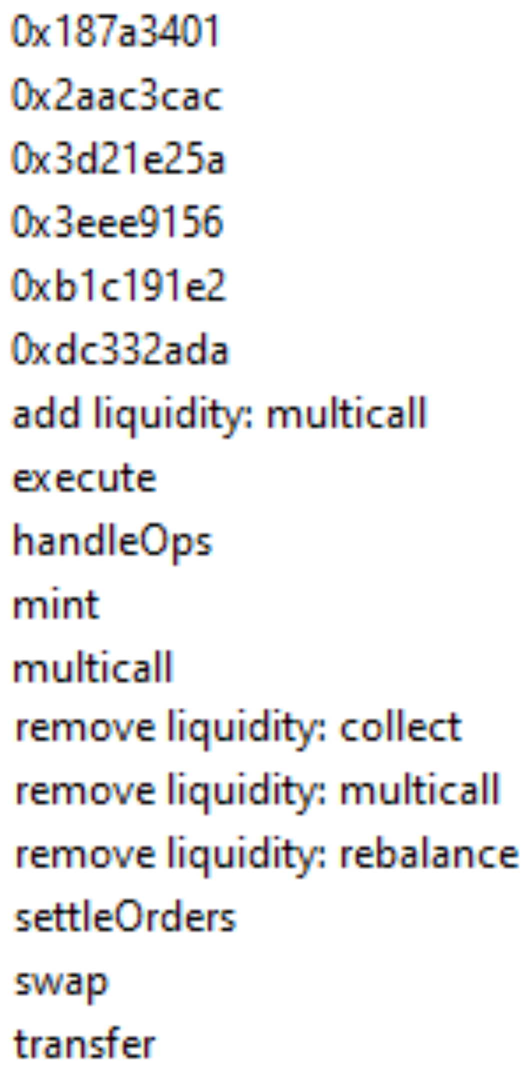
Autorité
des marchés
financiers

DeFi transactions

Different types of interactions

- Transfers
- Swaps
- Liquidity providing

But reality is more complex



0x187a3401
0x2aac3cac
0x3d21e25a
0x3eee9156
0xb1c191e2
0xdc332ada
add liquidity: multicall
execute
handleOps
mint
multicall
remove liquidity: collect
remove liquidity: multicall
remove liquidity: rebalance
settleOrders
swap
transfer

In practice: smart contracts

Token

- Different type ERC 20, 1155
- 3 basic parameters
- 6 basic functions:

Pool

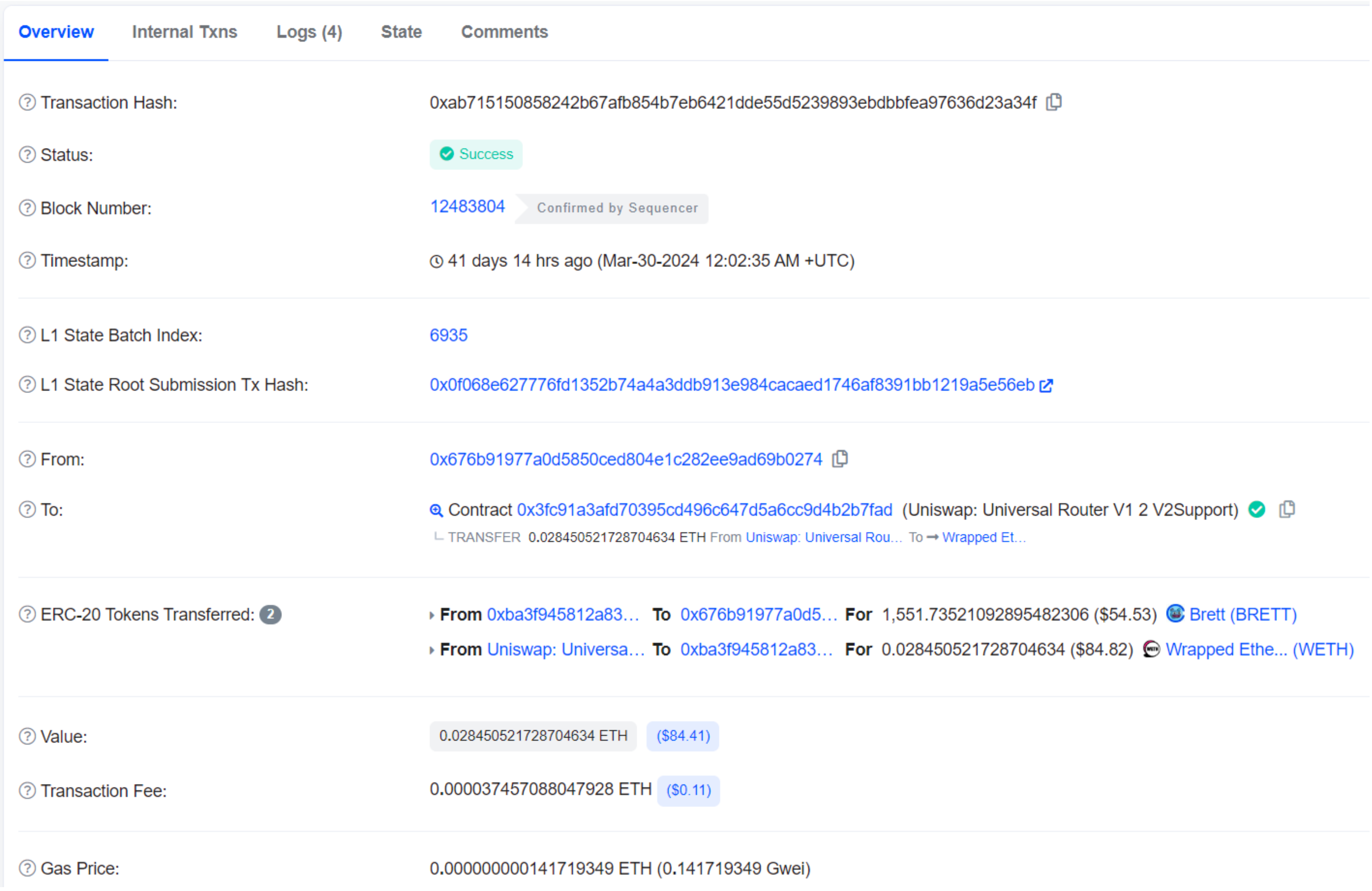
- Provides liquidity to DEX
- Allows to swap tokens
- Uses AMM to set price
- Reward liquidity providers

Router

- Find best path for the transaction
- Convert non ERC20 tokens to ERC20 tokens

Sources: https://basescan.org/

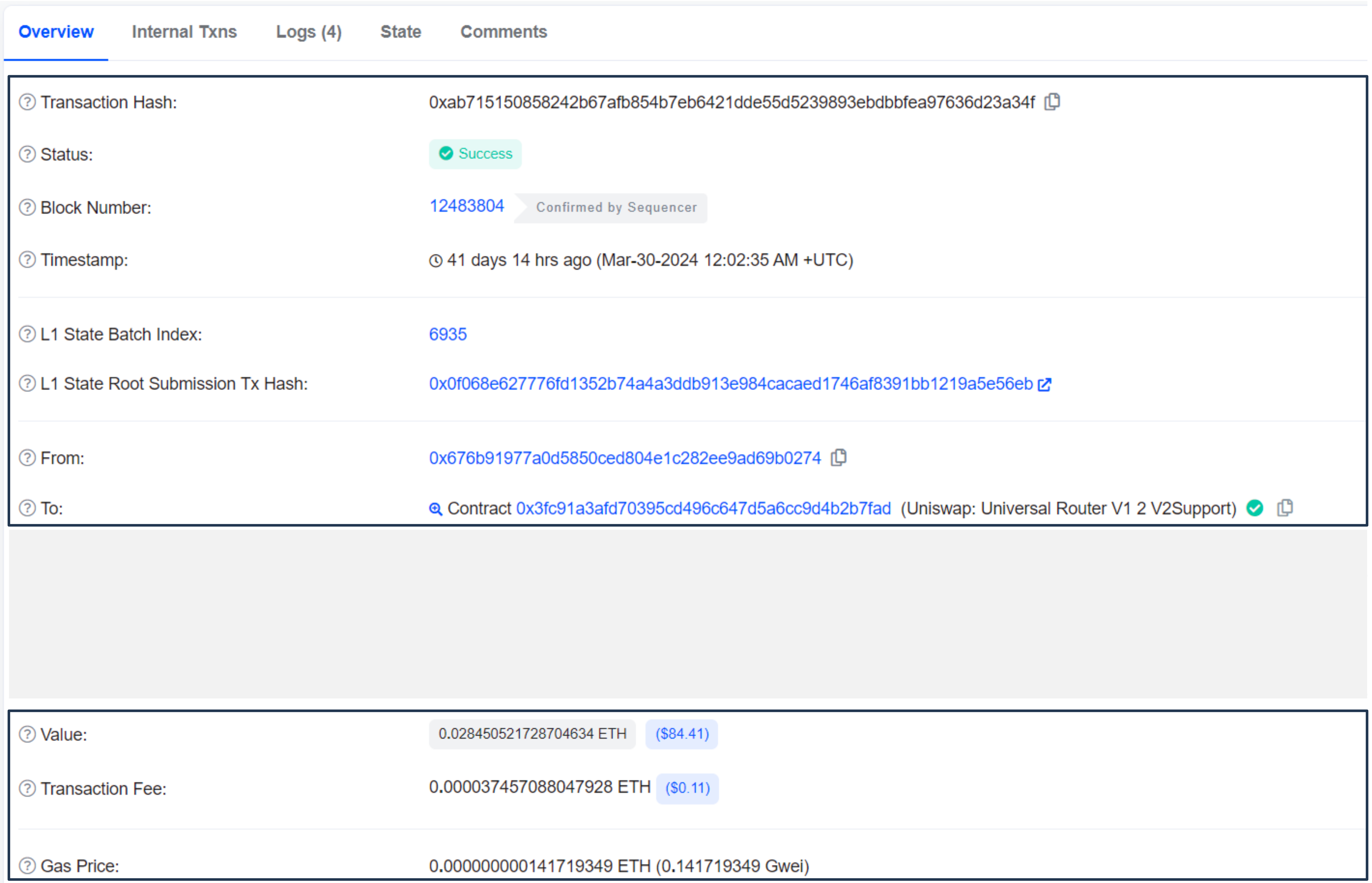
Examples of swap transactions



Transaction Hash:	0xab715150858242b67afb854b7eb6421dde55d5239893ebdbbfea97636d23a34f
Status:	✓ Success
Block Number:	12483804 Confirmed by Sequencer
Timestamp:	41 days 14 hrs ago (Mar-30-2024 12:02:35 AM +UTC)
L1 State Batch Index:	6935
L1 State Root Submission Tx Hash:	0x0f068e627776fd1352b74a4a3ddb913e984cacaed1746af8391bb1219a5e56eb
From:	0x676b91977a0d5850ced804e1c282ee9ad69b0274
To:	Contract 0x3fc91a3afd70395cd496c647d5a6cc9d4b2b7fad (Uniswap: Universal Router V1 2 V2Support) ✓ └ TRANSFER 0.028450521728704634 ETH From Uniswap: Universal Rou… To → Wrapped Et…
ERC-20 Tokens Transferred: 2	From 0xba3f945812a83… To 0x676b91977a0d5… For 1,551.73521092895482306 ($54.53) Brett (BRETT) From Uniswap: Universa… To 0xba3f945812a83… For 0.028450521728704634 ($84.82) Wrapped Ethe... (WETH)
Value:	0.028450521728704634 ETH ($84.41)
Transaction Fee:	0.000037457088047928 ETH ($0.11)
Gas Price:	0.000000000141719349 ETH (0.141719349 Gwei)

Sources: https://basescan.org/

Three in one: a simple one



Transaction Hash: 0xab715150858242b67afb854b7eb6421dde55d5239893ebdbbfea97636d23a34f

Status: Success

Block Number: 12483804 Confirmed by Sequencer

Timestamp: 41 days 14 hrs ago (Mar-30-2024 12:02:35 AM +UTC)

L1 State Batch Index: 6935

L1 State Root Submission Tx Hash: 0x0f068e627776fd1352b74a4a3ddb913e984cacaed1746af8391bb1219a5e56eb

From: 0x676b91977a0d5850ced804e1c282ee9ad69b0274

To: Contract 0x3fc91a3afd70395cd496c647d5a6cc9d4b2b7fad (Uniswap: Universal Router V1 2 V2Support) ✓

Value: 0.028450521728704634 ETH ($84.41)

Transaction Fee: 0.000037457088047928 ETH ($0.11)

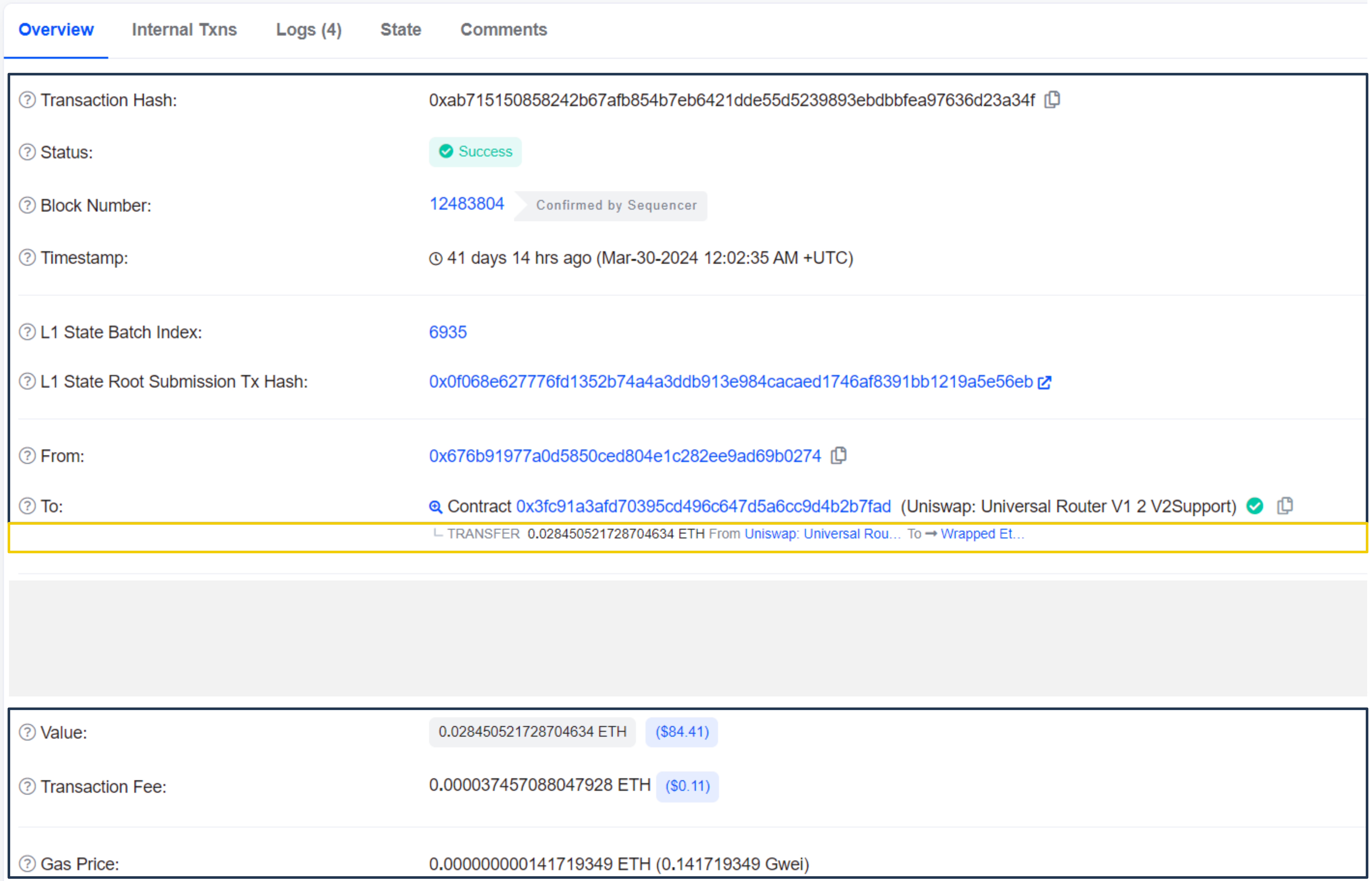
Gas Price: 0.000000000141719349 ETH (0.141719349 Gwei)

1. Main transaction

0,02845 ETH sent to router

Sources: https://basescan.org/

Three in one: a simple one



Transaction Hash:	0xab715150858242b67afb854b7eb6421dde55d5239893ebdbbfea97636d23a34f
Status:	Success
Block Number:	12483804 Confirmed by Sequencer
Timestamp:	41 days 14 hrs ago (Mar-30-2024 12:02:35 AM +UTC)
L1 State Batch Index:	6935
L1 State Root Submission Tx Hash:	0x0f068e627776fd1352b74a4a3ddb913e984cacaed1746af8391bb1219a5e56eb
From:	0x676b91977a0d5850ced804e1c282ee9ad69b0274
To:	Contract 0x3fc91a3afd70395cd496c647d5a6cc9d4b2b7fad (Uniswap: Universal Router V1 2 V2Support) ✓ └ TRANSFER 0.028450521728704634 ETH From Uniswap: Universal Rou... To → Wrapped Et...
Value:	0.028450521728704634 ETH ($84.41)
Transaction Fee:	0.000037457088047928 ETH ($0.11)
Gas Price:	0.000000000141719349 ETH (0.141719349 Gwei)

1. Main transaction

0,02845 ETH sent to router

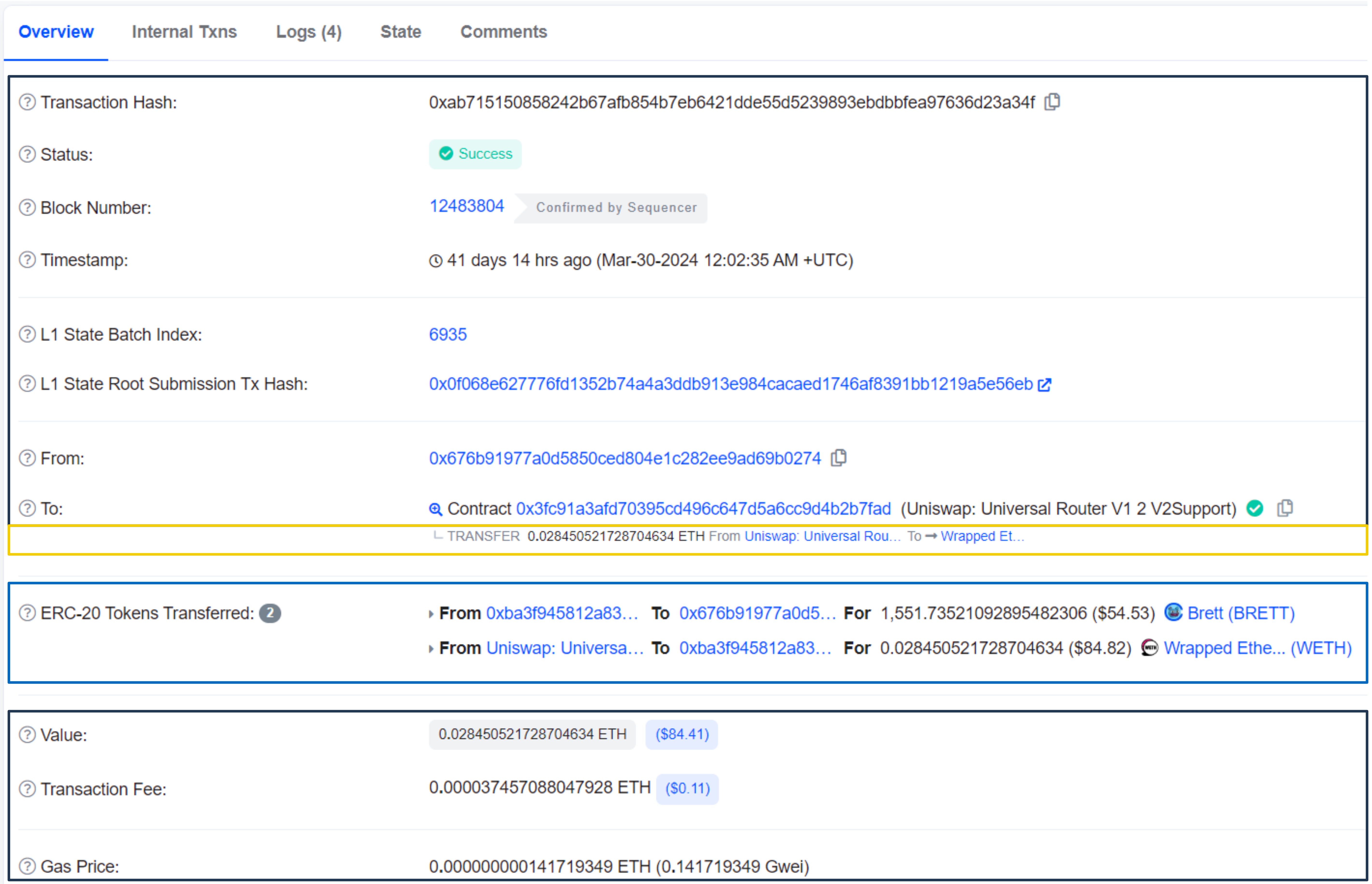
2. Internal transaction

ETH converted to WETH (ERC20 token)

Autorité des marchés financiers

Sources: https://basescan.org/

Three in one: a simple one



Transaction Hash:	0xab715150858242b67afb854b7eb6421dde55d5239893ebdbbfea97636d23a34f
Status:	Success
Block Number:	12483804 Confirmed by Sequencer
Timestamp:	41 days 14 hrs ago (Mar-30-2024 12:02:35 AM +UTC)
L1 State Batch Index:	6935
L1 State Root Submission Tx Hash:	0x0f068e627776fd1352b74a4a3ddb913e984cacaed1746af8391bb1219a5e56eb
From:	0x676b91977a0d5850ced804e1c282ee9ad69b0274
To:	Contract 0x3fc91a3afd70395cd496c647d5a6cc9d4b2b7fad (Uniswap: Universal Router V1 2 V2Support) ✓ └ TRANSFER 0.028450521728704634 ETH From Uniswap: Universal Rou... To → Wrapped Et...
ERC-20 Tokens Transferred: 2	▸ **From** 0xba3f945812a83... **To** 0x676b91977a0d5... **For** 1,551.73521092895482306 ($54.53) Brett (BRETT) ▸ **From** Uniswap: Universa... **To** 0xba3f945812a83... **For** 0.028450521728704634 ($84.82) Wrapped Ethe... (WETH)
Value:	0.028450521728704634 ETH ($84.41)
Transaction Fee:	0.000037457088047928 ETH ($0.11)
Gas Price:	0.000000000141719349 ETH (0.141719349 Gwei)

1. Main transaction

   0,02845 ETH sent to router

2. Internal transaction

   ETH converted to WETH (ERC20 token)

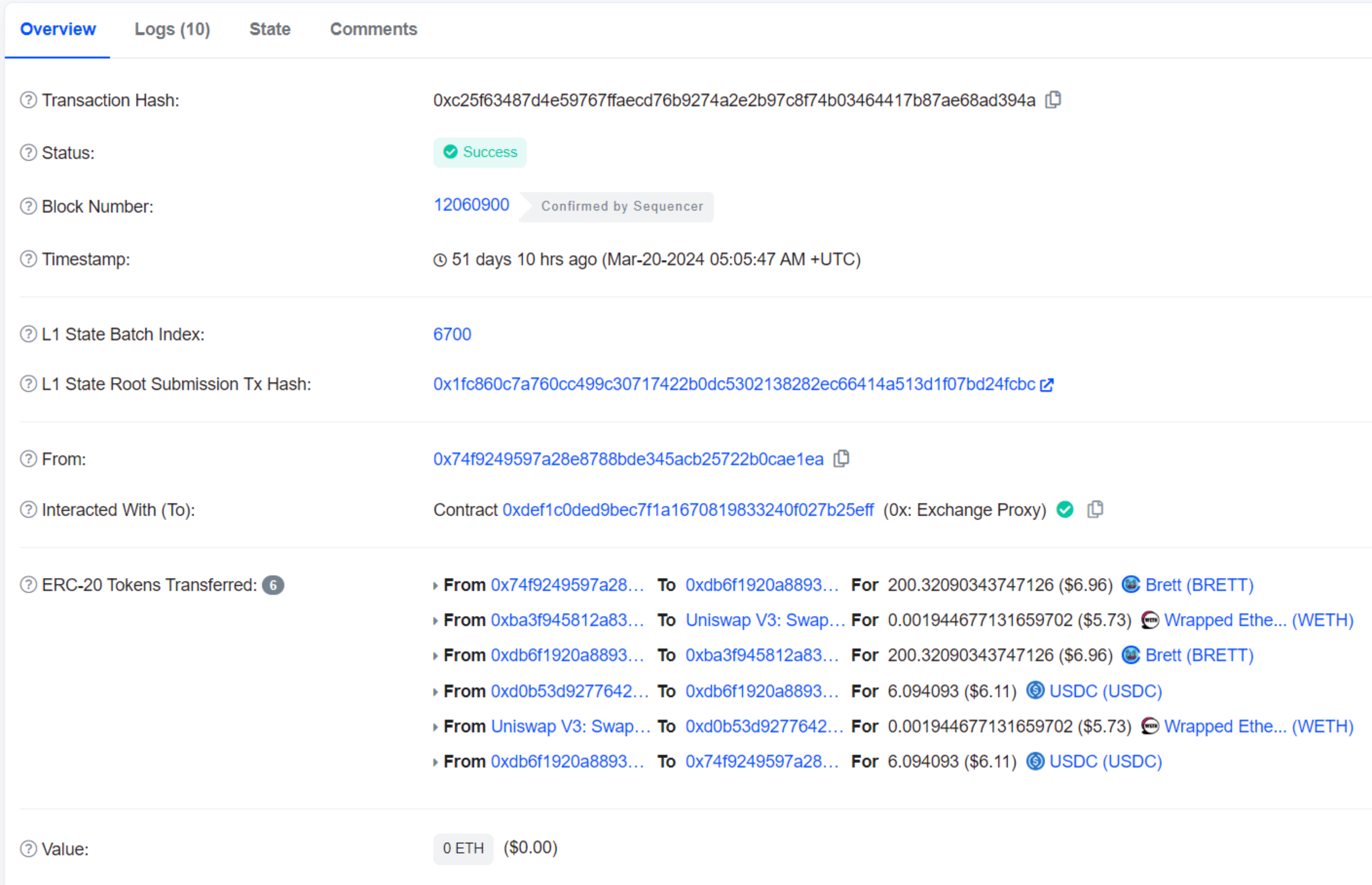
3. ERC transactions:

   0,02845 WETH swap for 1 551,74 BRETT

Autorité
des marchés
financiers

Sources: https://basescan.org/

Medium spicy



Overview | Logs (10) | State | Comments

Transaction Hash:	0xc25f63487d4e59767ffaecd76b9274a2e2b97c8f74b03464417b87ae68ad394a
Status:	Success
Block Number:	12060900 Confirmed by Sequencer
Timestamp:	51 days 10 hrs ago (Mar-20-2024 05:05:47 AM +UTC)
L1 State Batch Index:	6700
L1 State Root Submission Tx Hash:	0x1fc860c7a760cc499c30717422b0dc5302138282ec66414a513d1f07bd24fcbc
From:	0x74f9249597a28e8788bde345acb25722b0cae1ea
Interacted With (To):	Contract 0xdef1c0ded9bec7f1a1670819833240f027b25eff (0x: Exchange Proxy)
ERC-20 Tokens Transferred: 6	**From** 0x74f9249597a28… **To** 0xdb6f1920a8893… **For** 200.32090343747126 ($6.96) Brett (BRETT) **From** 0xba3f945812a83… **To** Uniswap V3: Swap… **For** 0.001944677131659702 ($5.73) Wrapped Ethe... (WETH) **From** 0xdb6f1920a8893… **To** 0xba3f945812a83… **For** 200.32090343747126 ($6.96) Brett (BRETT) **From** 0xd0b53d9277642… **To** 0xdb6f1920a8893… **For** 6.094093 ($6.11) USDC (USDC) **From** Uniswap V3: Swap… **To** 0xd0b53d9277642… **For** 0.001944677131659702 ($5.73) Wrapped Ethe... (WETH) **From** 0xdb6f1920a8893… **To** 0x74f9249597a28… **For** 6.094093 ($6.11) USDC (USDC)
Value:	0 ETH ($0.00)

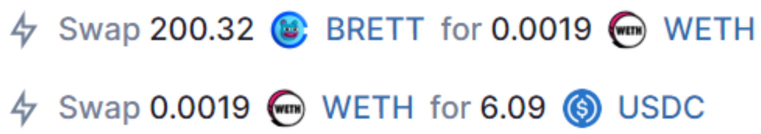
1. Main transaction

   nothing

2. Internal transaction

   nothing

3. ERC transactions:



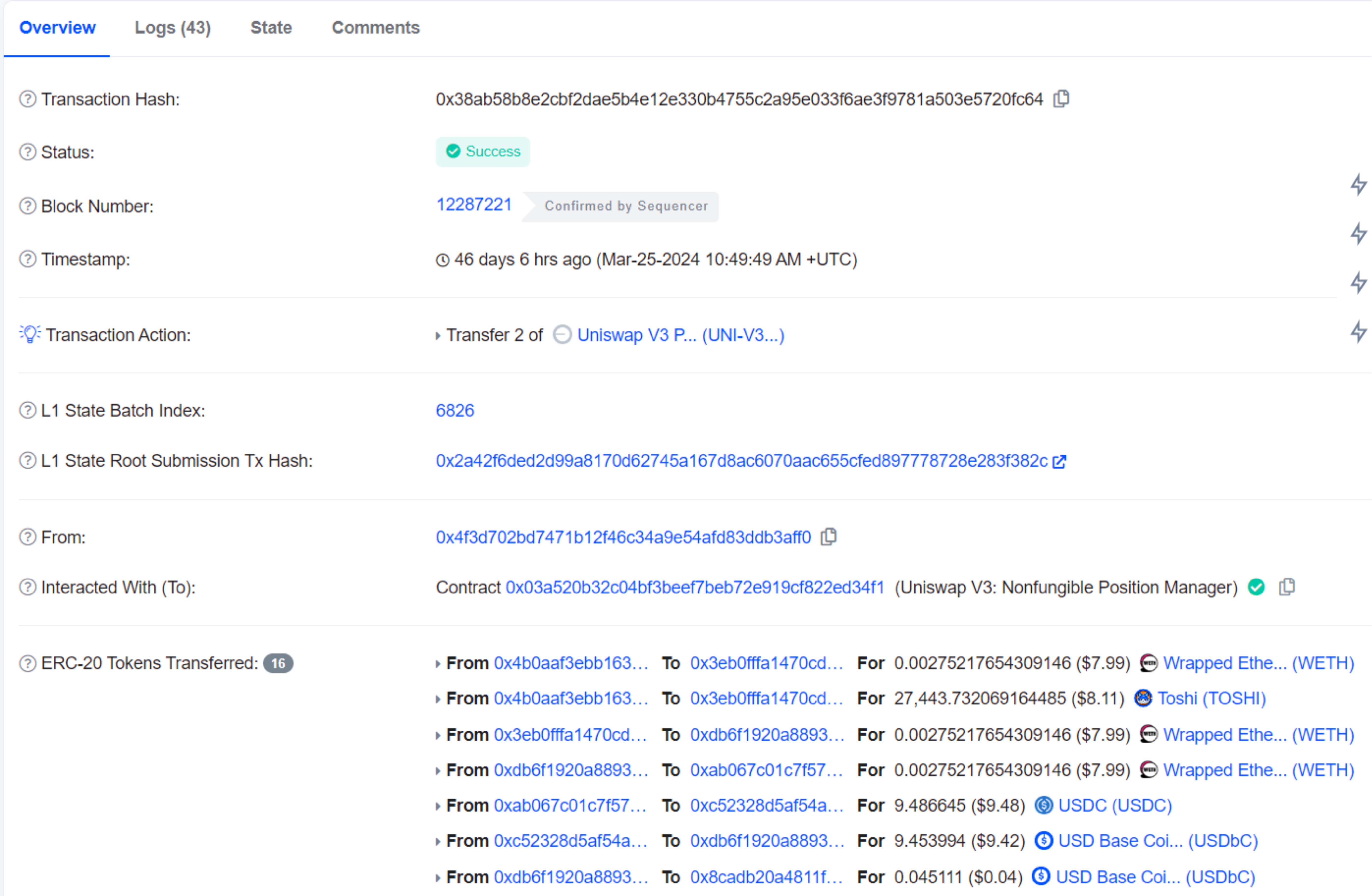
Swap 200.32 BRETT for 0.0019 WETH

Swap 0.0019 WETH for 6.09 USDC

Autorité des marchés financiers

Sources: https://basescan.org/

A complex case



Overview | Logs (43) | State | Comments

Transaction Hash: 0x38ab58b8e2cbf2dae5b4e12e330b4755c2a95e033f6ae3f9781a503e5720fc64

Status: Success

Block Number: 12287221 Confirmed by Sequencer

Timestamp: 46 days 6 hrs ago (Mar-25-2024 10:49:49 AM +UTC)

Transaction Action: Transfer 2 of Uniswap V3 P... (UNI-V3...)

L1 State Batch Index: 6826

L1 State Root Submission Tx Hash: 0x2a42f6ded2d99a8170d62745a167d8ac6070aac655cfed897778728e283f382c

From: 0x4f3d702bd7471b12f46c34a9e54afd83ddb3aff0

Interacted With (To): Contract 0x03a520b32c04bf3beef7beb72e919cf822ed34f1 (Uniswap V3: Nonfungible Position Manager) ✓

ERC-20 Tokens Transferred: 16

- From 0x4b0aaf3ebb163… To 0x3eb0fffa1470cd… For 0.00275217654309146 ($7.99) Wrapped Ethe... (WETH)
- From 0x4b0aaf3ebb163… To 0x3eb0fffa1470cd… For 27,443.732069164485 ($8.11) Toshi (TOSHI)
- From 0x3eb0fffa1470cd… To 0xdb6f1920a8893… For 0.00275217654309146 ($7.99) Wrapped Ethe... (WETH)
- From 0xdb6f1920a8893… To 0xab067c01c7f57… For 0.00275217654309146 ($7.99) Wrapped Ethe... (WETH)
- From 0xab067c01c7f57… To 0xc52328d5af54a… For 9.486645 ($9.48) USDC (USDC)
- From 0xc52328d5af54a… To 0xdb6f1920a8893… For 9.453994 ($9.42) USD Base Coi... (USDbC)
- From 0xdb6f1920a8893… To 0x8cadb20a4811f… For 0.045111 ($0.04) USD Base Coi... (USDbC)

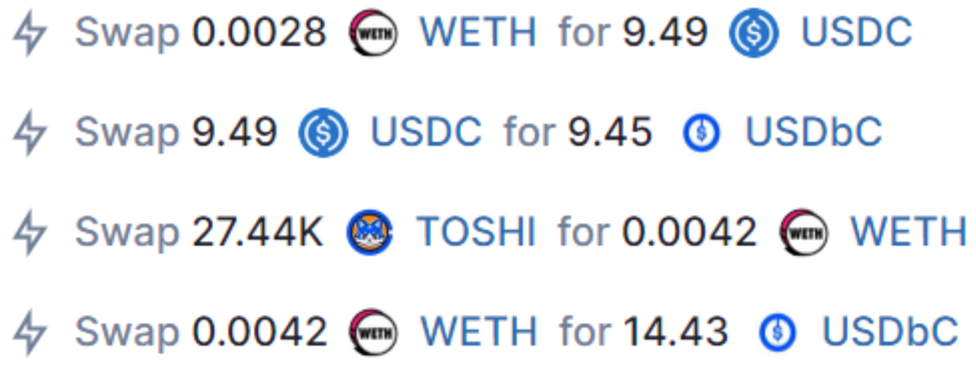
1. Main transaction

   nothing

2. Internal transaction

   nothing

3. ERC transactions:



Swap 0.0028 WETH for 9.49 USDC

Swap 9.49 USDC for 9.45 USDbC

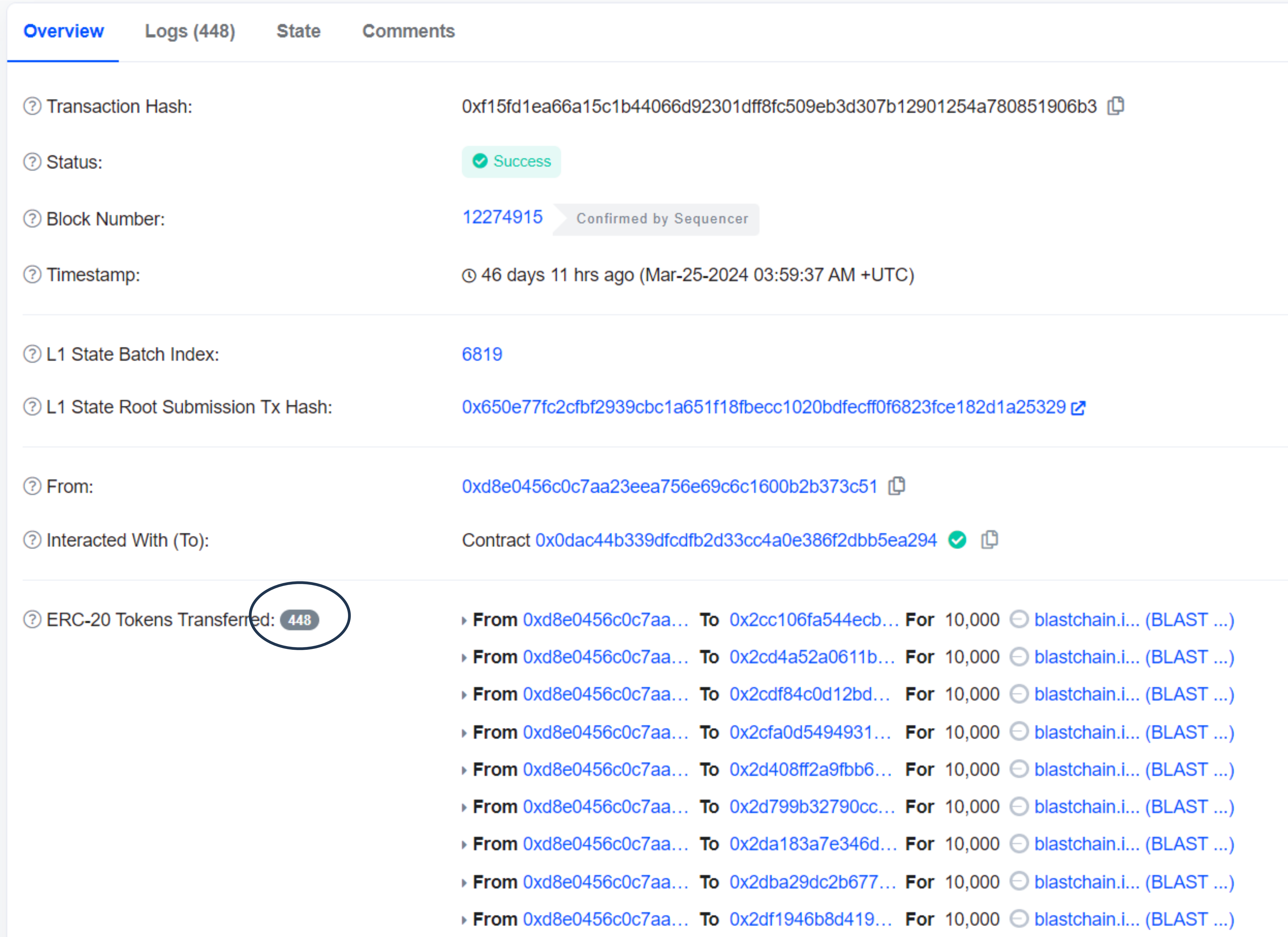
Swap 27.44K TOSHI for 0.0042 WETH

Swap 0.0042 WETH for 14.43 USDbC

Autorité des marchés financiers

Sources: https://basescan.org/

The exotic one



Transaction Hash: 0xf15fd1ea66a15c1b44066d92301dff8fc509eb3d307b12901254a780851906b3

Status: Success

Block Number: 12274915 Confirmed by Sequencer

Timestamp: 46 days 11 hrs ago (Mar-25-2024 03:59:37 AM +UTC)

L1 State Batch Index: 6819

L1 State Root Submission Tx Hash: 0x650e77fc2cfbf2939cbc1a651f18fbecc1020bdfecff0f6823fce182d1a25329

From: 0xd8e0456c0c7aa23eea756e69c6c1600b2b373c51

Interacted With (To): Contract 0x0dac44b339dfcdfb2d33cc4a0e386f2dbb5ea294

ERC-20 Tokens Transferred: 448

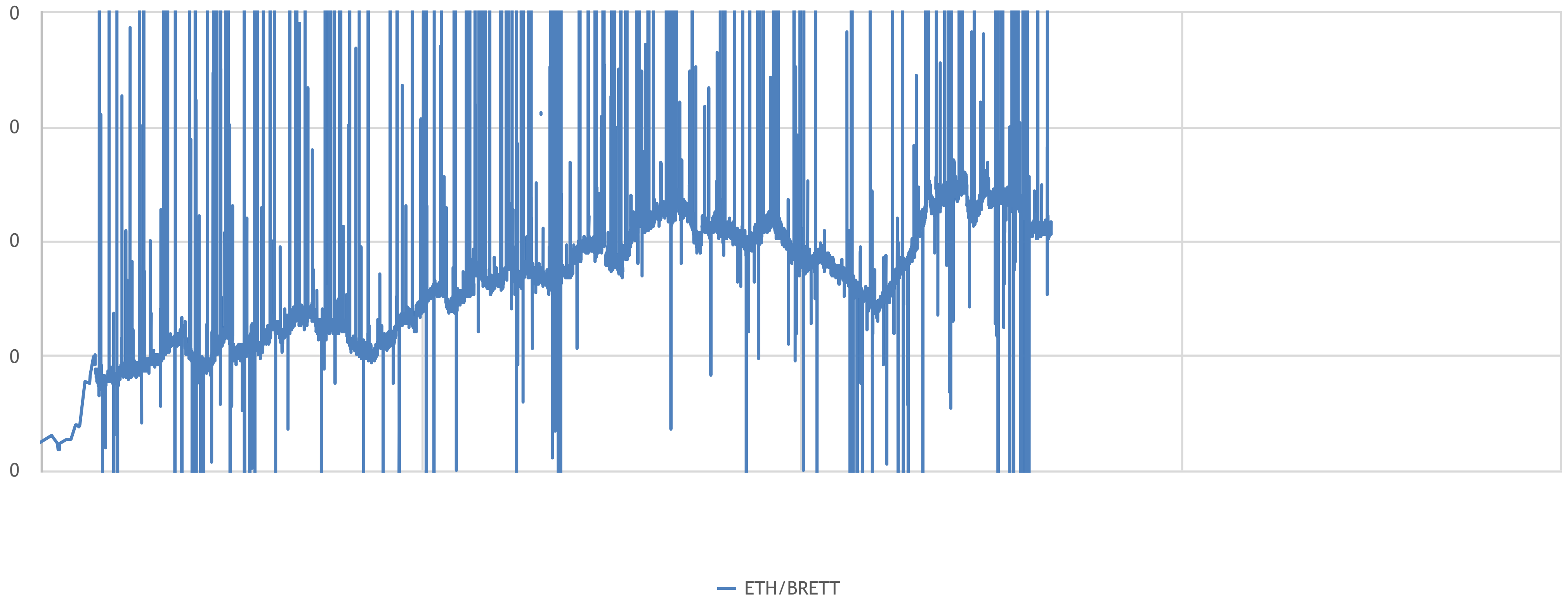
From 0xd8e0456c0c7aa… To 0x2cc106fa544ecb… For 10,000 blastchain.i... (BLAST ...)
From 0xd8e0456c0c7aa… To 0x2cd4a52a0611b… For 10,000 blastchain.i... (BLAST ...)
From 0xd8e0456c0c7aa… To 0x2cdf84c0d12bd… For 10,000 blastchain.i... (BLAST ...)
From 0xd8e0456c0c7aa… To 0x2cfa0d5494931… For 10,000 blastchain.i... (BLAST ...)
From 0xd8e0456c0c7aa… To 0x2d408ff2a9fbb6… For 10,000 blastchain.i... (BLAST ...)
From 0xd8e0456c0c7aa… To 0x2d799b32790cc… For 10,000 blastchain.i... (BLAST ...)
From 0xd8e0456c0c7aa… To 0x2da183a7e346d… For 10,000 blastchain.i... (BLAST ...)
From 0xd8e0456c0c7aa… To 0x2dba29dc2b677… For 10,000 blastchain.i... (BLAST ...)
From 0xd8e0456c0c7aa… To 0x2df1946b8d419… For 10,000 blastchain.i... (BLAST ...)

Autorité des marchés financiers

Our approach

Transaction extraction steps:

- Extract Main, Internal and ERC transactions for each transactions
- Trace each transaction using our algorithm (95% of the transactions can be traced)



Still have a large number of bad tracing

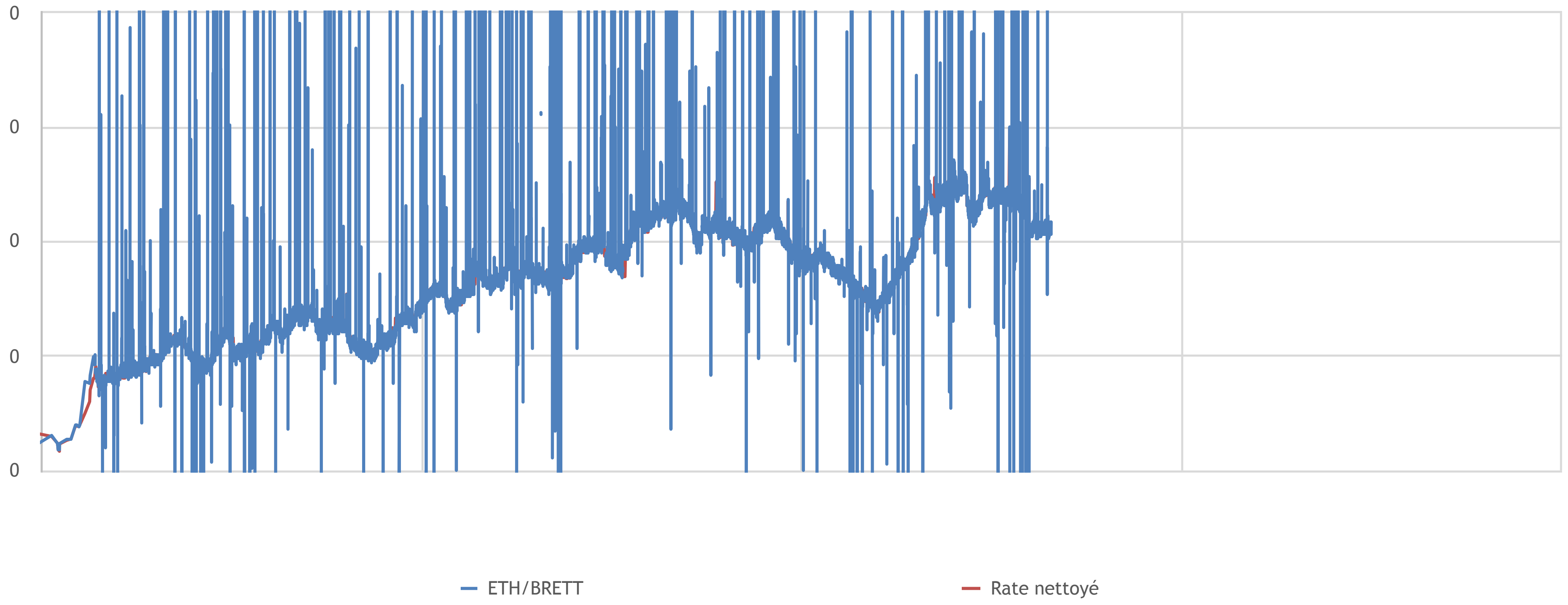
Autorité des marchés financiers

Our approach

Transaction extraction steps:

- Extract Main, Internal and ERC transactions for each transactions

- Trace each transaction using our algorithm (95% of the transactions can be traced)



- Data cleaning process to find theorical conversion rate

Autorité des marchés financiers

Our approach

Transaction extraction steps:

- Extract Main, Internal and ERC transactions for each transactions
- Trace each transaction using our algorithm (95% of the transactions can be traced)
- Extrapolate predicted rate and look for close ETH value in ERC transfers
- If found (5% variability is accepted), we replace value and complete tracing process

Finally

Pros

- Ability to reconstruct 99% of transactions
- Fast process if the transaction is simple
- Can be used to generate training datasets
- Independent from smart function code

Cons

- Slow process for complex transaction tracing
- Incomplete data

Finally

Pros

- Ability to reconstruct 99% of transactions
- Fast process if the transaction is simple
- Can be used to generate training datasets
- Independent from smart function code

Cons

- Slow process for complex transaction tracing
- Incomplete data

Objectives

1 – Use machine learning to increase speed and reliability

2 – Create a robust workflow to new dataset and adapt to futur updates and new technologies

Autorité
des marchés
financiers

Thank you

Questions?

LAMBDA
Laboratoire analytique, modélisation et BI de L'AUTORITÉ DES MARCHÉS FINANCIERS