Runtime Monitoring of Smart Contracts

On the Ethereum network

Lars Stegeman [s1346466]
l.stegeman@student.utwente.nl

April 18, 2018

Contents

1 Introduction **1**

2 Background **1**
2.1 The blockchain . . . 2
2.2 Smart Contracts . . . 2
2.3 Smart Contract exploits . . . 2
2.4 EVM . . . 2

3 Solidity **2**

4 Example **2**
4.1 Simple Token . . . 3
4.2 Specification . . . 3
4.3 Implementation of RuntimeMonitoredSimpleToken . . . 4
4.4 Discussion . . . 6

5 Runtime monitoring **6**

6 Property specification **6**

7 Related Work **6**
7.1 Smart Contract Verification . . . 6
7.1.1 Static Analysis Tools . . . 7
7.1.2 Formal Verification Tools . . . 7
7.2 Smart Contract Languages . . . 7
7.2.1 Bamboo . . . 7
7.2.2 Vyper . . . 7
7.3 Other related work . . . 8
7.3.1 ContractLARVA . . . 8
7.3.2 Ethereuem-runtime-verification . . . 8
7.3.3 The Hydra Project . . . 8

8 Planning **8**

1 Introduction

2 Background

This section will discuss the background information that will be built upon further in the document. First we will briefly discuss the important parts of the Ethereum Blockchain. Then we will discuss the Smart Contracts in more detail.

2.1 The blockchain

The Ethereum platform is built upon a distributed public ledger . On this ledger the cryptocurrency ether is stored. It is opposed to Bitcoin based on an account based system and not unspend transaction output. There are two types of accounts, one is a default account in which a user controls the spending of funds through its private keys. These accounts are called "Externally owned Accounts". The other option is a "Contract Account", which means that it is managed by code only. The code is set when the contract is constructed and initialised on the blockchain. Contract accounts only execute code when they are called from other contracts. Each contract has a persistent storage which is also maintained on the blockchain. This means that the Ethereum blockchain consists of two parts. The first part is the transaction history and the other part is the storage of all the deployed smart contracts combined. Transactions are the only entity that make changes to the storage. At an higher level overview we could see the Ethereum network as a large state machine in which changes to the state are controlled by transactions. Transactions are grouped in blocks and these blocks are distributed over the network and validated by each node.

2.2 Smart Contracts

Smart Contracts on the Ethereum network consist of two parts. Each contract has a set of functions and a storage. The contract set of functions is defined by the contract code that is deployed with the contract. This contract code is EVM bytecode and is usually compiled from a higher level programming language. When the contract is created the storage is initially empty. Only the contract code can make changes and add data to the presistent storage, within this storage the state of the contract is maintained. Each new function call has an empty memory, this can also be used to store data. But this data is not presistent through transactions, it is only persistent within the transaction. There are also so called "logs", this storage can only be used to store data and not retrieve. This storage is usually used to provide data for the external world because it can be searched efficiently.
Functions are only executed when they are called by external contracts. For example if a fund is to be released after a certain amount of time (block number higher then a certain amount). These funds will not be automatically transferred once the time treshold is reached, they will only be released when the function is called again.

2.3 Smart Contract exploits

2.4 EVM

3 Solidity

The most used language to develop contracts on Ethereum is Solidity. In this section we will formally define the syntax and semantics of the language.

4 Example

In this section a small example will be introduced. This example helps understand the problem this research tries to solve. The tool that will parse specification is not implemented yet, that is why in this example the specification is translated by hand. First we will introduce the contract and what is it supposed to do. Next we will add the specifications to the contract that we wish to check for during runtime. After that we will look at the code of the runtime monitored contract, which is made by translating the specified properties to Solidity code and adding it to the existing contract. Lastly we will discuss what this means for the contract in terms of gas execution and additional changes that had to be made to the contract.

4.1 Simple Token

In this example we will use the contract SimpleToken. The contract code can be found on the Ethereum Foundation website [1]. It models a minimum viable token, to keep the state of the contract a `mapping` is used that maps `address` to `uint`. This mapping is kept in the contracts internal storage and is stored on the blockchain. It indicates which addresses holds how much of this contracts token and changes each time the `transfer` function is called. The `transfer` function requires two parameters an address (`_to`) and an uint(`_value`) which specifies the amount to be sent. The from address is determined from the global variable `msg` which is present in each transaction. When the contract is created the constructor will be called. In this constructor the initialSupply is given as a parameter. All the initial supply is given to the contract creator (`msg.sender`). The totalSupply value is assigned and cannot be changed after initialisation.
Note that this contract is not ERC20 compliant. ERC20 is the interface that most tokens use to implement the desired functionality. This interface is defined in order for all wallets and exchanges to be able to handle different tokens [2]. The main difference is that this contract does not have an `approve` mapping which lets users approve a certain transfer of tokens. Also this SimpleToken does not allow `minting` or `burning` of tokens, in other words the total supply is fixed. Below we can see the Solidity source code of the contract SimpleToken.

```
pragma solidity ^0.4.20;

contract SimpleToken {
  /* This creates an array with all balances */
  mapping (address => uint256) public balanceOf;
  /* TotalSupply is fixed for this token, and does not change. */
  /* It is assigned in the constructor */
  uint256 totalSupply;

  /* Initializes contract with initial supply tokens to the creator of the contract
       */
  function SimpleToken(uint256 initialSupply) public {
    // Give the creator all initial tokens
    balanceOf[msg.sender] = initialSupply;
    totalSupply = initialSupply;
  }

  /* Send coins */
  function transfer(address _to, uint256 _value) public {
    // Check if the sender has enough
    require(balanceOf[msg.sender] >= _value);
    // Check for overflows
    require(balanceOf[_to] + _value >= balanceOf[_to]);
    // Subtract from the sender
    balanceOf[msg.sender] -= _value;
    // Add the same to the recipient
    balanceOf[_to] += _value;
  }
}
```

4.2 Specification

The above section describes the implementation of the token contract. However there is also a specification given in words as to what the contract should do. A few properties of this specification can be declared explicitly using pre and postconditions or invariants. These properties are important to the functionality of the contract. Properties could be translated to corresponding Solidity code and added to the contract. This code is executed within each transaction and thus the properties are checked and validated at runtime.
The first property is that the value of `totalSupply` is always equal to the sum of all the individual balances in the mapping `balanceOf`. Another property is that when a transfer function is executed the balance of the `_to` address is incremented with the `_value`. And the balance of the sender is decreased with the same value. The first property can be seen as an invariant of the contract and should hold before and after execution of a function. The second property should be checked after the execution of the transfer function and can be represented as a postcondition on that function.

[1] `https://www.ethereum.org/token`
[2] `https://github.com/ethereum/EIPs/blob/master/EIPS/eip-20.md`

The exact syntax on how to declare these properties is not defined yet. This example will use syntax that is close to the syntax that is used by JML.

```
pragma solidity ^0.4.20;

contract SimpleToken {
  /* This creates an array with all balances */
  mapping (address => uint256) public balanceOf;
  /* TotalSupply is fixed for this token, and does not change. */
  /* It is assigned in the constructor */
  uint256 totalSupply;


  /* @invariant totalSupply ==  sum = (forall x in balanceOf [ sum +=balanceOf[x]])

  /* Send coins */
  /* @ensures
      balanceOf[_to] == \old(balanceOf[_to]) + _value &&
      balanceOf[msg.sender] == \old(balanceOf[msg.sender]) - _value &&
      forall x : x != _to || x != msg.sender : balanceOf[x] == \old(balanceOf[x])
  */
  function transfer(address _to, uint256 _value) public {
  }
}
```

4.3 Implementation of RuntimeMonitoredSimpleToken

The specification together with the solidity code should be compiled to a new smart contract. The compiled smart contract will have the same behaviour but with extra assertions added to the code. The tool should be able to parse specifications and add the correct code to the corresponding functions. For this example the code will be made by hand because the tool still needs to be developed. The contract solidity code for the Runtime-Monitored SimpleToken can be seen in the snippet below.
For this example if an assertion in the specification is false, the call will return an error. Later more complex behaviour can be added when this situation occurs. The way the contract works from the outside should not change, because several front-end implementations could depend on it. This means that the added behaviour should be inside the functions that are in the original contract. To accomplish this the transfer function body is moved to a seperate private function `transfer_body()`. The original transfer function will do several things. First it stores the current state in the `memory` storage of the contract. This is needed because specifications can use the `\old` keyword to reference to variables before the function execution. Next it checks the invariant and possible preconditions. After that the function body is executed. Lastly the postcondition and invariant are checked, the state before execution is passed as a parameter to the functions.

```
pragma solidity ^0.4.20;

contract RuntimeMonitoredSimpleToken {
  /* This creates an array with all balances */
  mapping (address => uint256) public balanceOf;
  /* TotalSupply is fixed for this token. */
  /* It is assigned in the constructor */
  uint256 totalSupply;


  /* ------------------------------------------------------------ */
  /* Array to keep a list of in use addresses */
  address[] public addressesInUse;
  /* Struct to save the state of the balanceOf mapping */
  struct BalanceOfStruct{
    address _address;
    uint256 balance;
  }
  /* ------------------------------------------------------------ */


  /* Initializes contract with initial supply tokens to the creator of the contract
      */
  function RuntimeMonitoredSimpleToken(uint256 initialSupply) public {
    // Give the creator all initial tokens
```

```
    balanceOf[msg.sender] = initialSupply;
    /* ------------------------------------------------------------ */
    addressesInUse.push(msg.sender);
    /* ------------------------------------------------------------ */
    totalSupply = initialSupply;
}

/* Send coins */
function transfer(address _to, uint256 _value) public {
    //Save state of current variables needed for validation of properties
    BalanceOfStruct[] memory  old_balanceOf = new BalanceOfStruct[](addressesInUse.
        length);
    address[] memory old_addressesInUse = new address[](addressesInUse.length);
    for(uint x; x < addressesInUse.length; x++){
      old_addressesInUse[x] = addressesInUse[x];
      old_balanceOf[x] = BalanceOfStruct(addressesInUse[x], balanceOf[
          addressesInUse[x]]);
    }

    // check invariant
    invariant();
    // check requires

    // execute body
    transfer_body(_to,_value);
    // check ensures
    transfer_ensures(_to, _value, old_addressesInUse, old_balanceOf);
    // check invariant
    invariant();
}

function transfer_ensures(address _to, uint256 _value, address[]
    _old_addressesInUse, BalanceOfStruct[] _old_balanceOf ) private{
    uint index_to;
    uint index_from;
    for(uint x; x< addressesInUse.length; x++){
      if(addressesInUse[x] == _to){
        index_to=x;
      }
      if(addressesInUse[x] == msg.sender){
        index_from = x;
      }
    }
    uint old_balance_from = _old_balanceOf[index_from].balance;
    uint old_balance_to ;
    if(index_to >= _old_addressesInUse.length){
        old_balance_to = 0;
    }else{
        old_balance_to = _old_balanceOf[index_to].balance;
    }

    bool exp1 = (balanceOf[_to] == (old_balance_to + _value));
    bool exp2 = (balanceOf[msg.sender] == (old_balance_from - _value));
    bool exp3 = true;
    for(x=0; x < _old_addressesInUse.length && exp3; x++){
      if(x != index_to || x!= index_from){
        exp3 == (balanceOf[addressesInUse[x]] == _old_balanceOf[x].balance);
      }
    }
    assert(exp1 && exp2 && exp3);
}

function invariant() private{
    uint256 sum;
    for(uint x; x < addressesInUse.length; x++){
      sum += balanceOf[addressesInUse[x]];
    }
    assert(totalSupply == sum);
}

/* transfer function original body */
function transfer_body(address _to, uint256 _value) private {
    // Check if the sender has enough
```

```
    require(balanceOf[msg.sender] >= _value);
    // Check for overflows
    require(balanceOf[_to] + _value >= balanceOf[_to]);
    // Subtract from the sender
    balanceOf[msg.sender] -= _value;
    // Add the same to the recipient
    balanceOf[_to] += _value;
    /* ------------------------------------------------------------ */
    bool inUse = false;
    for(uint x; x < addressesInUse.length && !inUse; x++){
      if(addressesInUse[x] == _to){
          inUse = true;
      }
    }
    if(!inUse){
        addressesInUse.push(_to);
    }
    /* ------------------------------------------------------------ */
  }
}
```

4.4 Discussion

There are a few choices that had be made during the implementation of the runtime contract.
The first observation is that a `mapping` variable is not iterable. This means that properties that use quantifiers and reference mappings cannot be tested at runtime. Mappings in Solidity do not store a keyset and just store the information at the corresponding hash of the key. This limitation can be avoided if we store the used keyset in an additional array. This array must be kept in the storage since it must be peristent between transactions. The SimpleToken contract has only one mapping and the storage array `addressesInUse` keeps track of the keys. A more general approach would be to define a new storage construction called IterableMapping [3].
Another limitation is that mappings cannot be declared in `memory`. Thus to store the variable `\old(someMapping{ key => value}` there is need for either a additional mapping in `storage` or the mapping can be represented by a struct array (which can be stored in `memory`). The last approach is used in the example, but additional analysis is needed to determine which method is the most gas efficient. Gas cost is significantly higher for `storage` in comparison to `memory`, but additional actions are required to have the information converted to a struct variable.
Lastly because of additional checks the gas cost for each function is not constant anymore. This effect is worse compared to the increase of gas cost in normal contracts. For example if the SimpleToken contract had 1000 addresses in use the cost of each function call would not increase in the original contract. But in the runtime-monitored contract the gas costs would increase since the array `addressesInUse` has length 1000 which causes extra iterations for validation. The additional gas cost are not analysed for this SimpleToken contract, but in the final tool tests with multiple transactions should be analysed for their gas cost.

5 Runtime monitoring

6 Property specification

7 Related Work

There is a lot of work related to this topic. There are papers discussing the verification of smart contracts. They can be catagorised as static analysis or formal verification. Additionally other contract languages have been proposed to help writing secure smart contracts. The last subsection discusses some other related work.

7.1 Smart Contract Verification

Due to the recent exploits that were executed on the Ethereum blockchain this has sparked an interest in the formal verification area. There are many proposals of verification tools that will

[3] `https://github.com/ethereum/dapp-bin/blob/master/library/iterable_mapping.sol`

help to write secure smart contracts. Because if the bytecode of a contract is committed to the blockchain it cannot be changed afterwards. This means that testing and verification of the code before commiting it to the network is important. They can be categorized in two groups; static analysis and formal verification. The first class are tools that analyse the EVM code or a higher level code and check for patterns. Patterns that are known to be vulnerable get reported. The code is not actually executed, only symbolically. The second group is formal verification. These tools work by giving a specification and the program. And proof that the program is correct for all possible inputs with respect to the given specification. Note that the solidity code is usually translated to EVM or some intermediate language in which the proofs can be more easily automated.

7.1.1 Static Analysis Tools

There are many tools that are defined in this area. Most of the tools have the same functionality. You can analyse contracts, either by Solidity code or EVM bytecode. These contracts can be analysed locally or from an online provider (Ethereum mainnet or one of the test nets). Examples are Mythril [1], Securify [2] and Oyente [3]. The Oyente tool also offers the possibility to analyse all the contracts on the whole blockchain. Their tool is not only available on Github but also has a paper which describes the choices made fo the analysis tool.

7.1.2 Formal Verification Tools

Verification in this fields means that a specification has to be written. Specification gives meaning to what the contract should do, i.e. a contract should behave according to its specification. However because Solidity is not fit for this most tools are defined at the EVM bytecode level, or introduce an intermediate contract language. These programs are then proven correct considering all possible inputs with respect to the given specification. KEVM [4] and eth-isabelle [5] are very similar. Both tools are able to execute a large set of the official ethereum test suite and are able to proof specifications correct for certain contracts. Other approaches use an intermediate language over which properties can be proven correct. Lolisa [6] and Scilla [7] fall under this category.

7.2 Smart Contract Languages

Smart contracts are usually written in a high level language that compiles to EVM (Ethereum Virtual Machine) bytecode. Currently the best known and most used language is Solidity (as described in detail in section 3). But there are other options available that compile to the same EVM. They differ in their syntax and influences by other languages.

7.2.1 Bamboo

Bamboo is a morphing smart contract language. State transitions the core part of the language design. This makes the state transitions in smart contracts explicit. This way it avoids reentrancy by default. Instead of having a global state of the contract, contracts morph into a new contract by calling functions. This way there should be less suprises in the execution of smart contracts. The project is located in a repository located at `https://github.com/pirapira/bamboo`. As an example the smart contract for a crowd funding is used. The crowd funding usually has several stages in which different things can happen. In Solidity these stages are usually modeled using boolean variables and enforced using `modifiers`. With this approach it is hard to keep track which functions are enabled at which state. In Bamboo this is not the case since functions are declared within an state and functions modify the signature of the smart contract.

7.2.2 Vyper

Vyper is a new and experimental smart contract programming language. It is maintained by the Ethereum Foundation at `https://github.com/ethereum/vyper`. The idea is to limit certain functions and aspects that are possible in Solidity to make writing smart contracts less error prone. It also tries to make smart contracts more human readable to make it simpler to see what will happen when a function is called. For example `modifiers`, inline assembly and class inheritance is not allowed in Vyper as opposed to Solidity.

7.3 Other related work

A number of other proposals have been published which try to make smart contracts more secure. They do not belong to a certain category but are related to the current work. Some projects only have source code available and do not have documentation or a paper.

7.3.1 ContractLARVA

ContractLARVA can be found on github at `https://github.com/gordonpace/contractLarva`. Following the instructions on the README you can write a specification and a contract in Solidity. The compiler will combine these two and output a new Solidity contract with the runtime verification checks in place. Properties have to be specified using *dynamic event automata* (DEA) [8]. The tool is based on a similar tool called LARVA for Java.

7.3.2 Ethereuem-runtime-verification

This project is located at `https://github.com/shaunazzopardi/ethereuem-runtime-verification`. No documentation is available for this project. It mentions the LARVA project in the description in that it differs from LARVA because this runtime-verification tool can dynamically add properties to an already deployed smart contract.

7.3.3 The Hydra Project

The Hydra Framework is a project for smart contracts on the Ethereum network. It tries to make smart contracts more secure by making multiple implementations of the same contract. They call this *N-of-N-version programming*. The different implementations are controlled by a meta contract which forwards the incoming calls to all the implementations. If the implementations do not agree on a single answer, the meta contract will be able to react on this. When such a vulnerability is found a bounty is given to the person who exploited the vulnerability. They call this principle *the exploit gap*, this means that a hacker should claim the bounty instead of exploiting the vulnerability. More information can be found in their paper [9].

8 Planning

References

[1] "Mythril," https://github.com/ConsenSys/mythril.

[2] "Securify," https://securify.ch/.

[3] L. Luu, D.-H. Chu, H. Olickel, P. Saxena, and A. Hobor, "Making smart contracts smarter," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2016, pp. 254–269.

[4] E. Hildenbrandt, M. Saxena, X. Zhu, N. Rodrigues, P. Daian, D. Guth, and G. Rosu, "KEVM: A complete semantics of the Ethereum virtual machine," Tech. Rep., 2017.

[5] Y. Hirai, "Defining the ethereum virtual machine for interactive theorem provers," in *International Conference on Financial Cryptography and Data Security*. Springer, 2017, pp. 520–535.

[6] Z. Yang and H. Lei, "Lolisa: Formal syntax and semantics for a subset of the solidity programming language," *arXiv preprint arXiv:1803.09885*, 2018.

[7] I. Sergey, A. Kumar, and A. Hobor, "Scilla: a smart contract intermediate-level language," *arXiv preprint arXiv:1801.00687*, 2018.

[8] C. Colombo, G. J. Pace, and G. Schneider, "Dynamic event-based runtime monitoring of real-time and contextual properties," in *Formal Methods for Industrial Critical Systems (FMICS)*, ser. Lecture Notes in Computer Science, vol. 5596, L'Aquila, Italy, 2008, pp. 135–149.

[9] L. Breidenbach, P. Daian, F. Tramer, and A. Juels, "Enter the hydra: Towards principled bug bounties and exploit-resistant smart contracts."