WhitePaper: Cyber-Security Variant of TruCol protocol

Eliminating triage intermediaries for zero-day exploits using a decentralised payout protocol.

Chihab Amghane
Radboud University

Victoria Bosch
Radboud University

Rashim Charles
Radboud University

Marc Droogh
Delft University of Technology

Clara Main
Radboud University

Chinari Subhechha Subudhi
Delft University of Technology

Akke Toeter
Delft University of Technology, Radboud University

Eric van der Toorn
Delft University of Technology

1 Introduction

This document presents a Trustless Security protocol that aims to help ethical hackers retrieve their bounties without ambiguity, whilst simultaneously enabling companies to show their customers how much money is staked on their open source software stacks being secure against zero-day exploits.

To explain how the protocol may help both of these stakeholders (ethical hackers and companies using open source software), we will first describe, what we think is, a typical procedure for vulnerability disclosures in Section 2. Then we will explain how the protocol can improve on that in Section 3. Next, Section 4 describes strategies to specify how the protocol may be implemented. The limitations and weaknesses of our strategy and protocol are detailed in Section 5. This white-paper is concluded in Section 8.

2 Assumptions

This sections present some of the assumptions that are made about the current way zero-day exploits, and vulnerabilities in general are treated.

2.1 Ethical Hacker Perspective

1. We assume it is not always as easy and/or attractive for whitehat hackers/ethical hackers to publish an exploit and retrieve an accompanying financial reward for the publication. This assumption is based on popular media such as darknet diaries, posts on `news.ycombinator.com`, communications with two ethical hackers and possibly other sources. This assumption is based on (a combination of) the following sub-assumptions:

a. Vulnerabilities may be discovered at small/non-profit software development companies that have not allocated a large budget fraction to security.
b. Ambiguity in the specification of the bug bounty/reward program may be interpreted in the advantage of the company during triage.
c. The triage process may take a relatively long time, requiring the ethical hacker to have sufficient funds to sustain living costs coverage until the pay-out.
d. A conservative/carefulness in the ethical hacker towards approaching the company with respect to the legality of discovering the vulnerability may hinder/slow down the vulnerability disclosure process.
e. The effort required to contact the company and convince them of the seriousness of the bug may consume unnecessary resources.

2.2 Company - perspective

1. We assume cybersecurity vulnerabilities become increasingly more relevant in our increasingly more digitized world. This assumption may be seen as being substantiated by for example the *Cyber Security Assessment Netherlands 2021 (CSAN 2021)* as presented by the Dutch National Coordinator Counterterrorism and Safety of the Ministry of Justice and Security. Currently, there is only the Dutch version available at: `https://www.nctv.nl/onderwerpen/cybersecuritybeeld-nederland/documenten/publicaties/2021/06/28/cybersecuritybeeld-nederland-2021`. We assume that this trend can be extrapolated from a Dutch perspective to a more global perspective, given the international media coverage of many ransomware attacks.
2. We assume that companies are interested, or will become more interested, in showing their customers and/or stakeholders (a quantified perspective on) *how* secure their technology is. We assume it can be quite challenging to convey this perspective clearly due to the following factors:
   a. Vulnerabilities can be found in various sections of the company, ranging from social engineering, misconfiguration to zero-day exploits. It is difficult to give customers a comprehensive yet concise/simple insight in how "secure" all these attack surfaces are.
   b. The impact of a vulnerability may be ambiguous or not easily quantifiable. For example, for some companies, vulnerabilities may allow malicious actors to take over critical infrastructure, whilst other vulnerabilities may lead to data leaks or other undesired side effects.
   c. It may be difficult to accurately assess the capabilities of malicious adversaries.
3. We assume some companies might be unfamiliar with vulnerability disclosure and accompanying triage processes. These delicate processes may seem intimidating for new companies that want to start paying attention to their cybersecurity, and this may lead to a lower allocation of cybersecurity budget.

3 Protocol

This section presents the protocol and explains how it can improve the way vulnerability disclosures are completed for deterministically verifiable zero-day exploits.

3.1 Scope

The protocol can be applied to identify more vulnerabilities than deterministically verifiable zero-day exploits. Companies can also for the decentralised Virtual Machine and add a

specific configuration, and add a bounty on that forked decentralised VM. This way, a hacker may leverage the particular configuration to find an exploit. This procedure also allows the protocol to identify some supply-chain attack vulnerabilities. For example, if an invalid certificate is used to compromise the device.

However, both misconfiguration and supply chain attack partially deviate from the main benefit of collective nature of the protocol. For example, it may incentivise hackers to focus efforts on particular configurations, that are not (necessarily) useful for other companies. However, at the same time, hackers could still opt to focus on the mutual elements of all forked decentralised virtual machines to collect the bounties with a single, more powerful exploit. This scope/applicability of the protocol is visualised in Figure 1.

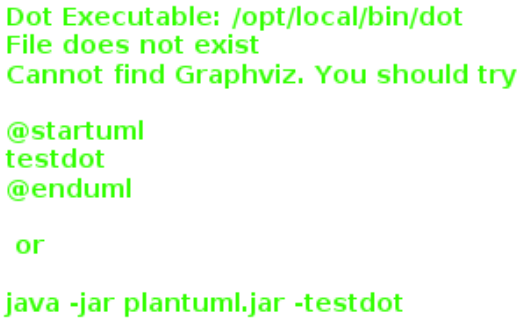

Figure 1 The proposed TruSec protocol is not suited to deal with social engineering attacks, nor is it ideal for misconfiguration exploits and/or supply-chain attacks. Instead, it is designed to increase the rate of discovery of deterministically verifiable zero-day exploits. Note, we acknowledge that attacks can be, and often are, a combination of the types.

With respect to Figure 1, the following notes are made:

1. The orange attack types imply the proposed protocol is not designed to tackle these issues, nor does it provide full coverage (against malicious agents) for these attack types. However:
   a. The misconfiguration could be covered if companies upload their configurations into DVMs. These configurations would typically not benefit from the collaborative staking, as it is less likely that other companies happen to use the same configurations.
   b. Some of the supply chain attacks could be covered if the ethical hackers are able to propagate these supply chain exploits into the DVMs.

3.2 Usage

With this scope defined, one can look at how companies and ethical hackers interact according to the proposed protocol.

The basic idea is that companies and users (stakeholders) can put their open source software stacks on a decentralised virtual machine (DVM). They can then collectively stake money on the security of the stacks, such that everyone can see how much money says: *the use of certain software packages/combinations is safe.* This enables companies, to show their customers for example:

With us, your data is stored using MongoDB Version 5.1, $314.159,- says it is uncompromised, and it's running on Ubuntu Server version 21.10, which has $4.200.000,- staked on its security. This setup has a configuration with a security on which we staked $9001,-. If

any of these software packages get compromised by whitehat hackers, we will be the first to know.

We believe that might be clear language that enables decision makers and customers interested in company A, to get an intuitive understanding on *how secure* some (critical) segments of the company A software are.

For the whitehat/ethical hackers, the advantages are clear; they know before they start their work how large their payout will be, and they get a direct payout upon completion (after the predetermined responsible disclosure period has ended).

3.2.1 Disclaimer

The presented protocol does not provide a insight in the complete security of a system/company. As visualised in Figure 1, the protocol does not cover all attack surfaces of companies. Hence, if other attack surfaces, such as social engineering are used, companies can still get compromised, regardless of the amount they staked. Therefore, it is important that the numerical value of the amount staked on the zero-day exploit security level is not abused to convey a false sense of security by the staking companies to their customers.

3.3 Description

The protocol is shown in Figure 2.

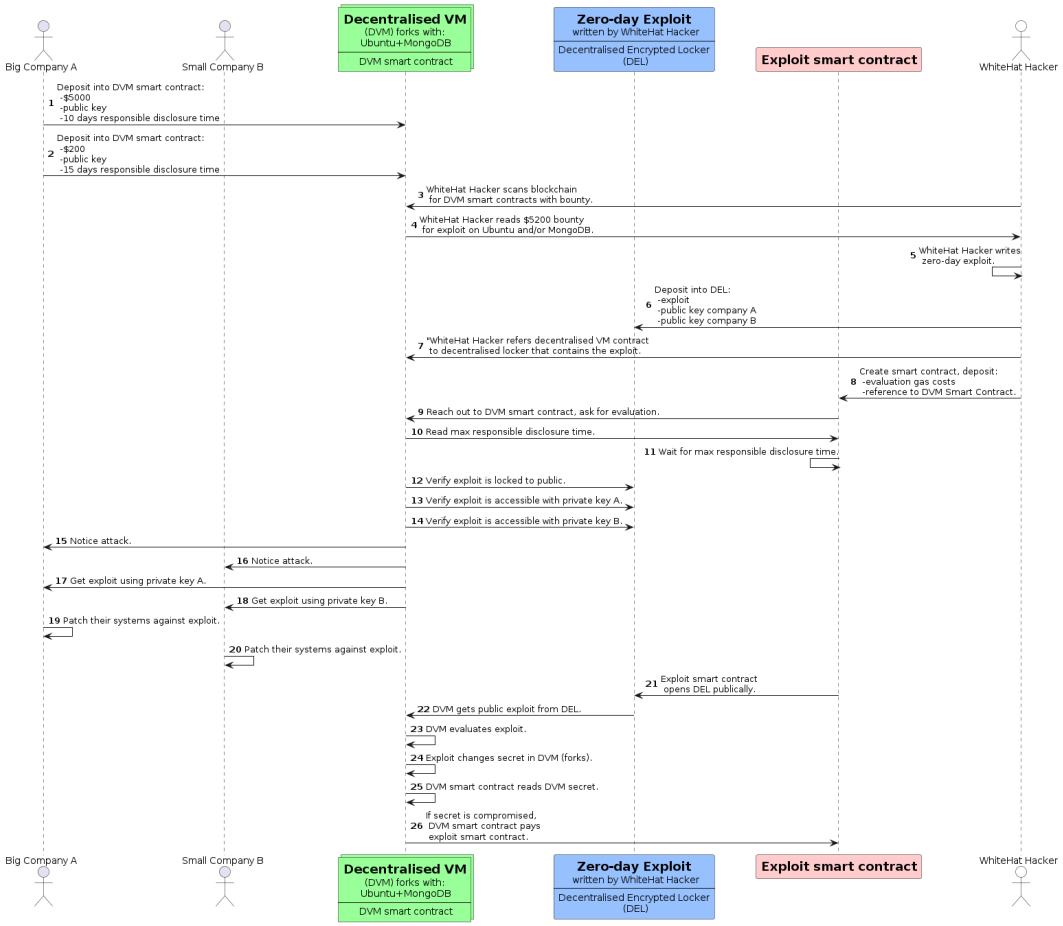

Figure 2 Visualisation of the interaction of the TruSec protocol. This is an ever-lasting cycle, where at the end of the process, companies can re-deploy the patched decentralised stack, and allocate new funds. Whitehat hackers can scan for new attacks.

3.4 Figure 2 notes

With respect to Figure 2, the following notes are made:

1. The attack written by the ethical hacker should be accessible on chain, such that everyone can verify that the attack indeed compromises the decentralised VM/honeypot. This is critical for the automatic payout.
2. The decentralised locker is used to prevent malicious hackers to inspect/copy the attack before the responsible disclosure period is over.
3. It would be better if the contract specifies the locker location, allowing the staking companies to actively check if an attack is found, instead of the attack reaching out to the honeypot. This is because the latter could attract unwanted attention. However, these are currently considered implementation details.

4 Implementation

Implementation details are omitted at this stage. One can note that developing decentralised virtual machines for security purposes requires a significant effort, even when considering

ports from e.g. the Ethereum Virtual Machine (EVM).

5 Discussion

The presented proposal for protocol development can be critically evaluated. This section aims to identify possible weak points.

5.1 Limitations

The following limitations are identified in the protocol:

1. The proposed protocol, in its initial form, does not (necessarily) work for security compromises that are not clearly pre-defined. For example, if the decentralised virtual machine/stack/honeypot is configured to only pay-out in case an internal value/secret is modified, a whitehat hacker might be able to gain read-access to the secret, which could be considered a hack, but the whitehat hacker would not receive a payout. Accordingly, companies may specify different payouts to different types of security breaches. This may reduce the added value of collaborative staking.
2. The costs of running a decentralised virtual machine, along with their interactions, are currently expected to be relatively high based on e.g. the costs of roughly 50 dollars for a single Ethereum transaction.
3. We expect most hacks do not rely on pure zero-day exploits, accordingly we think the scope of this protocol is significantly limited w.r.t. the complete cybersecurity threat landscape.
4. This protocol will most likely not allow companies to test their entire system, as we currently consider it practically infeasible to simulate the various types of social engineering and or interactions with non-decentralised platforms (on a blockchain). So companies cannot make claims about their overall level of cybersecurity based on this protocol alone.
5. This protocol does not protect against economically irrational malicious agents. Examples could be:
   a. Actors with revenge sentiment. They could for example skip the payout and use zero-day exploits to hurt a company that staked their open source software stack.
   b. Nation states may not care about payouts and instead use found zero-day exploits themselves, instead of disclosing them.

5.2 Related Work

It was noted during the TechEx conference, that companies like Google and Microsoft already fund vulnerability disclosures for, for example, Ubuntu. This can be seen as collective funding, hence one could argue the added value of the proposed protocol may be limited in this respect.

Additionally, there are companies like HackerOne that perform independent triage, hence one could argue the added value of doing this in a decentralised fashion is limited.

6 Questions

Your advice and expert knowledge within the domain of cybersecurity and ethical hacking is asked in particular on the following questions:

1. Would you perhaps be able to give us an approximation on the V, W, X, Y, Z percentages of cyberattacks as displayed in Figure 1?

   a. Note, we will have to do our own due-diligence on this, however, as a first indicator/ballpark estimate, your perspective would significantly move us forward in assessing the (potential) real-world impact of the proposed protocol.
   b. Are there relevant attack strategies that we omitted? (That are perhaps (indirectly) suitable for the proposed protocol).
2. Based on your experience, would you expect the protocol to be of value in real-world applications?
   a. (If not), which bottlenecks do you identify?
3. Would you consider a talk at Blackhat feasible (assuming the work is done well), with as topic: a presentation of the protocol (specification) with/without working implementation?
   a. Note, understand you do not have an crystal ball, however, perhaps your team has more experience into typical Blackhat submission topics and trends, than us.
4. Are there perspectives that you would like to share? Did we miss any angles/relevant factors? Is there any advice you could give us, or research-directions that may be relevant in this endeavour?

7 Planning

This section summarises the information required to generate a planning towards a Call for Papers submission for the *Policy Track* of the Blackhat 2022 USA submission.

7.1 Schedule

The 2022 Blackhat (USA) edition takes place on 2022-08-06 to 2022-08-11. Even though the call for papers is not yet open, one can develop a planning analogue to the call for papers for the 2021 Blackhat USA edition. For that conference, the dates were specified as:
Source: `https://www.blackhat.com/us-21/call-for-papers.html`

- Call for Papers Opened: February 2, 2021
- Call for Papers Closed: April 10, 2021
- Notification to Submitters: end of May, 2021
- Event Dates: July 31 - August 5, 2021

Hence, shifting the planning with one week, since the 2022 edition will occur one week later:

- Call for Papers Opened: February 9, 2021
- Call for Papers Closed: April 17, 2021
- Notification to Submitters: end of May, 2021
- Event Dates: August 06 - August 11, 2022

7.2 Deliverables

To create a successful submission to the Blackhat 2022 (USA) edition, the following deliverables are required:

1. A track specification.
   - Source: `https://www.blackhat.com/html/tracks.html`
2. *Assumed:* Abstract specification
   - Source: `https://i.blackhat.com/docs/cfp-sample-submissions.pdf`
3. *Assumed:* Presentation Outline
   - Source: `https://i.blackhat.com/docs/cfp-sample-submissions.pdf`
4. *Assumed:* Attendee Takeaways
   - Source: `https://i.blackhat.com/docs/cfp-sample-submissions.pdf`

5. *Assumed:* Why Black Hat motivation.
   - Source: `https://i.blackhat.com/docs/cfp-sample-submissions.pdf`
6. *Assumed:* Presentation slides.
   - Source: Imagination.

7.2.1 Submission Requirements (ASIA)

The following Blackhat submission requirements are specified for the ASIA event:
Source: `https://www.blackhat.com/call-for-papers.html`

1. Submissions may only be entered by researchers/speakers (no submissions from PR firms/marketing representatives).
2. Black Hat does not accept product or vendor-related pitches. Black Hat will disqualify any product or vendor pitch.
3. Submissions must clearly detail the concepts, ideas, findings, and solutions a researcher or speaking team plans to present.
4. Submissions that highlight new research, tools, vulnerabilities, etc. will be given priority.
5. Submissions that include White Papers are highly encouraged and will also be given priority.
6. Black Hat will disqualify incomplete submissions; complete your submission in its entirety.
7. Individuals may submit more than one proposal, but each proposal must be submitted via a separate submission form.
8. Each submission must include detailed biographies of the proposed speaking team.
9. Submitters will be contacted directly if Review Board members have any questions about a submission.

7.2.2 Tailoring Submission

Suggested resources to tailor the submission to maximise acceptance probability:
Source: the recommendation section of: `https://asia-briefings-cfp.blackhat.com/`.

- Example submissions: `https://i.blackhat.com/docs/cfp-sample-submissions.pdf`
- Tips: `https://insinuator.net/2017/04/some-quick-tips-for-submitting-a-talk-to-black-hat-or-`
- Acceptance: `https://www.helpnetsecurity.com/2016/03/30/how-to-get-your-talk-accepted-at-bla`
- Tips: `https://hexsec.blogspot.com/2012/12/create-good-security-cfp-responses.html`
- Pitfall avoidance: `https://research.kudelskisecurity.com/2020/04/02/5-common-cfp-submission-mis`

8 Conclusion and Recommendations

The proposed protocol may enable companies to convey the level of security of (segments of) their open source technology stack more intuitively to their customers/stakeholders. Additionally, the protocol enables whitehat/ethical hackers to retrieve payouts directly without ambiguity.

The development of the protocol requires significant work, and it is currently not clear what the added value of the protocol would be in real-life settings. A presentation of the protocol, without working implementation, in the *Policy Track* of the Blackhat (USA) 2022 edition may be a direct probe to the cybersecurity world to assess the interest in actually implementing the protocol.