TRITON/TRISIS attack

Julia Vister

September 2025

Contents

Chapter 1

Introduction to TRITON/TRISIS

Triton, also known as Trisis and HatMan, is a malware discovered in August 2017 when petrochemical facilities in the Middle East was targeted. The malware attacked safety instrumented systems (SIS), which is a set of hardware or software controls that provides the last protection layer against a physical incident. SIS does an immediate shutdown on a critical system, or parts of it, if a dangerous condition is detected, and therefore protects human life. [9] The malware targeted the Schneider Electric's Safety Instrumented System, where it affected the Triconex safety controllers. Triton malware compromises these controllers by reprogramming them with custom payloads during execution. It consists of two main components: a PC-based module that communicates with the safety controller, and a malicious binary that is deployed directly onto the controller. [5] It is believed that the malware reads/writes programs, functions and the query state of a SIS controller. [6] [10]

There has not been a lot of progress in investigating the attack, other than diagnosing the attack right after the attack in 2017. In recent years, investigations have revealed that the attackers maintained a presence within the plants system for several years before launching the attack. This prolonged access indicate that a deep understanding of the ICS was essential. Attacks against SIS in particular demand a high level of process knowledge. The attackers spent this time studying stolen engineering docs like technical manuals, and by mapping out the network through enumeration. [10] It is also believed that they gained access to the network through spear phishing, but no evidence is found of this. [9] FireEye, a cybersecurity company, found an IP address in a left behind file that was registered to the Central Scientific Research Institute of Chemistry and Mechanics in Moscow.[11] Triton is attributed to TsNIIKhM as a Russian government-controlled research institution under Russia's Ministry of Defense.[6] They are also known as Temp.Veles and XENOTIME. [8] [3] This

institution supports the Russian army with cyber capabilities and research. [6] Triton has only been publicly observed in the 2017 attack and there are no variants publicly known.

The goal of the attacker was a remote takeover and shutting down the plant's SIS.[11] Fortunately, there was a flaw in the code, and another safety system noticed the change in code so the attacks did not succeed. No harm was therefore done in this attack, other than bringing the plant to a halt from shutdown. [9] If the attackers would have succeeded, achieving the goal of shutting down the SIS, the plant's last line of defense would have been gone. The next step of the attackers is believed to be to take over the mechanical systems, so they would then have access to run unsafe processes to do a lot of damage. Worst case scenario if the attack succeeded is the release of toxic hydrogen sulfide gas and result in a massive explosion, which would again lead to filling the air in the surrounding areas with toxic chemicals. [11]

Chapter 2

Evidence and Incidents

2.0.1 Official and Systematic Reports

- **NCSC (2017):** The UK National Cyber Security Centre published one of the first official advisories on TRITON. This report contributed to that the malware targeted Schneider Electric Triconex Safety Instrumented Systems (SIS) and highlighted the risks to operational technology. [10]
- **CISA (2019):** The U.S. Cybersecurity and Infrastructure Security Agency released Malware Analysis Reports (MAR-17-352-01, "HatMan"), and contributed to a technical breakdown of TRITON's components and how it works with SIS controllers. [5]
- **IC3/FBI (2022):** A joint advisory from the Department of Justice and FBI stated that TRITON remains an active threat to critical infrastructure and contributed to the fact that nation-state actors, in this case Russia, are likely behind its development and deployment. [6]
- **MITRE ATTACK (2024):** Description of the attack and attackers.[3][8]
- **MIT Technology Review (2019):** Although less technical, this report emphasized TRITON's significance as one of the first pieces of malware designed specifically to target human safety. [7]
- **Trellix (2020):** Contributed to analysis of how TRITON fits into a broader aspect of ICS-targeting malware and discussed its sophistication and persistence. [9]
- **Purdue University (2024):** A more recent academic perspective on the attack. Summarizes the attack, attribution, and consequences in the context of critical infrastructure protection. [11]

2.0.2 Documented Incidents

To date, the Triton malware has been publicly linked to a single confirmed incident, namely the attack on the petrochemical facility in Saudi Arabia in 2017. The US and UK government advisories attribute the malware to Russian state-linked actors associated with TsNIIKhM. [1] [2]
In 2021, the United States indicated a TsNIIKhM employee for attempting intrusions against U.S Energy Sector organizations, where the employee accessed the systems and deployed Triton. [1]

2.0.3 Industrial Incidents

The publicly confirmed industrial incident involving TRITON occurred in 2017 a the petrochemical facility. Attackers compromised Schneider Electric Triconex Safety Instrumented Systems. TRITON bypassed traditional IT defenses and attacked the systems responsible for maintaining safety in industrial operations. The malware triggered a fail-safe shutdown, halting production, but preventing more damage. Had the attack succeeded, it could have silently disabled the SIS. This would have led to consequences such as major gas explosions and loss of human life. Therefore, this incident is highly significant from an OT perspective. [9]

Chapter 3

Technical Analysis

3.0.1 Cyber Kill Chain

TRITON required long reconnaissance to gather information about the system which included gaining knowledge of Triconex, engineering documents and network topology. Next an offline weaponization step that is necessary for the next step to make sure payload is compatible and stealthy, and therefore a ready-to-deploy controller payload. Delivery into the enterprise (the plant) through an entry vector that is believed to be spear-phishing and stolen credentials. Here the attackers has gained an initial foothold that is user-level and is needed for further steps by placing tools where the software runs. The attacker has therefore established a presence. Exploitation of engineering tools and credentials to reach SIS, which is needed run loader against controllers and the attackers can issue legitimate looking operations. Installation of the malware where attackers write/upload capability to controller and staged payload on controller to its ready to be activated against SIS - this maintains control for the next step. In the Command and Control step the attackers move across IT to OT by using remote tools like WMImplant, RDP jump boxes and backdoor. [3] Attacker has sustained remote control and access to OT accounts which make them capable of monitoring SIS changes. Finally reprogramming SIS logic and sends unauthorized command messages to the Triconex safety controllers. Attackers have direct control over SIS behavior and trip a controller into a failed safe state that led to automatic shutdown of the plant. [3] See figure 3.1.

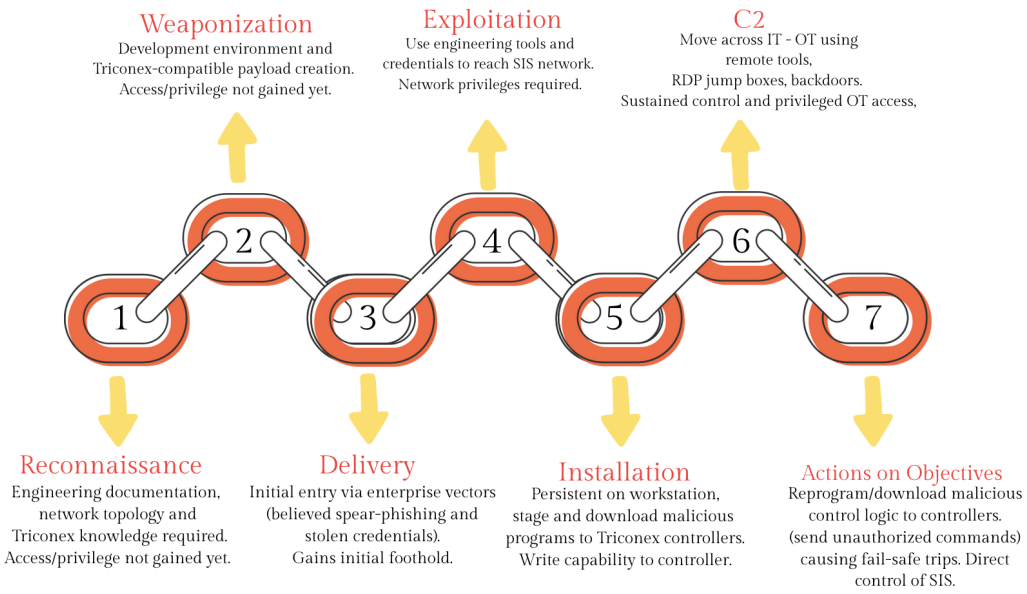

Figure 3.1: Cyber Kill Chain of TRITON

3.0.2 OSI 7-Layer Model

TRITON activity is documented from Physical upwards. Reports provide evidence of how the malware operated across Layers 2-7, with the data link, network, session and application layers being especially important as this is where reconnaissance, movement and payload deployment occurred.

- Layer 1: Physical Triconex safety controllers require a physical key switch to load new logic. This key switch has RUN and PROGRAM mode, where malicious downloads was possible if in PROGRAM mode. Some mitigation against this are to restrict console access and escort physical maintenance. and have the key switch at RUN.
- Layer 2: Data Link Attackers traversed poorly segmented networks and gained access to the SIS VLAN from compromised workstations. Mitigations include isolated SIS networks, log SIS connections and disabling unused ports.
- Layer 3: Network TRITON needed network routing from IT into OT. MITRE documents that the attackers used RDP(Remote Desktop Proto-

col) jump boxes to pivot into SIS networks.[3] Mitigation is to apply strict firewall rules and restrict traffic to authorized hosts.

- Layer 4: Transport TRITON payloads communicated through TriStation protocol, and the attackers used remote protocols such as RDP and WMI. While not a big attack vector, this layer provided the channels over which malicious payloads and control messages were carried and therefore plays a big supporting role. Mitigation is block unauthorized encrypted tunnels and deploy ICS-aware IDS to inspect traffic.

- Layer 5: Session Attackers maintained authenticated sessions to control workstations and issue SIS commands. Mitigations include limit privileged session duration and requiring MFA for remote sessions.

- Layer 6: Presentation Malicious content was packaged in formats recognizable by Triconex controllers as payloads were downloaded as controller programs. Rather than being the main focus, this layer supported the attacker greatly. [5] Mitigations include verifying integrity before downloads and using hashed controller logic files.

- Layer 7: Application TRITON exploited the TriStation software on Windows workstations to send malicious logic to SIS controllers. Mitigations include vendor patching, role-based access and dual approval for SIS changes.

Chapter 4

Mitigation Steps: Defense in Depth

If one layer fails, other layers still protect the system. The mitigation steps proposed are mapped to the Cyber Kill Chain stages identified in Section 3.

4.0.1 Reconnaissance

- People: Train staff not to share plant details online.
- Process: Enforce strict access controls, classify engineering documents as sensitive.
- Technology: Limit exposure of sensitive documentation and monitor for unusual network scans.

4.0.2 Weaponization

- People: Stay updated on adversary techniques.
- Process: Track actor TTPs.
- Technology: Detect known tool signatures.

4.0.3 Delivery

- People: Staff trained on recognizing phishing attempts, and reporting it.
- Process: Security awareness campaigns and phishing simulations.
- Technology: Email filtering, MFA.

4.0.4 Exploitation

- People: Trained engineers to validate SIS changes or alerts.
- Process: Dual-control approval before SIS changes, monitored sessions.
- Technology: Role-based access control, strict firewall rules between IT and OT.

4.0.5 Installation

- People: Trained operators to recognize anomalies such as unexpected trips and modified logic.
- Process: Frequent integrity checks.
- Technology: Endpoint detection and response on workstations, hash verification of controller logic.

4.0.6 Command and Control

- People: Review remote access logs reviewed by admin and investigation of alerts.
- Process: Session logging and approval.
- Technology: Network segmentation, monitoring RDP traffic.

4.0.7 Actions on Objectives

- People: Engineers and operators trained to verify and response to SIS changes.
- Process: Incident drills.
- Technology: Controller keyswitch in RUN, automatic alerts on logic changes.

Human error and awareness were crucial. Spear-phishing and credential theft targeted people directly, and weak processes allowed attackers to move further. A nation-state actor such as those behind TRITON has resources to remain stealthy for years before the attack, meaning pure technical defenses were not enough.

Chapter 5

Would it be the same today?

If the TRITON campaign was attempted today, modern AI and large language models could reduce the time needed for key stages of the attack, as well as reduce the expertise needed. AI tools could automatically scan open-source intelligence, generate diagrams and simulate engineering documents to speed up the time it takes to get familiar with ICS environments. Deepfake audio/video could also make spear-phishing campaigns more convincing and therefore easier to trick staff. Large language models can also assist attackers in writing and testing malicious payloads, reducing time needed in this step. AI could also streamline the adaptation of the malware to different versions or ICS protocols. [4]

In this case, defenders could also benefit from the use of AI and LLM's to improve resilience. LLM's can assist incident response by summarizing logs, correlating alerts and suggesting actions in real time. It could also simulate attack scenarios to test resilience and train staff to raise human awareness. AI can also help in prevention with AI-powered email filtering which can block phishing attempts, and automated privilege monitoring can flag suspicious use before SIS is reached.[4]

Chapter 6

ChatGPT

6.0.1 Question 3 and 4 using ChatGPT

See attachment for ChatGPT's results after generating the answers to question 3 and 4.

I agree with the general mapping, as both mine and ChatGPT's analysis agree that TRITON required long reconnaissance, and further validate the next steps in the cyber kill chain. However, there is a lack of specificity in tools that are documented in official reports. My version is more fact-specific.
When it comes to the OSI model, again I agree with the general mapping across it as it aligns with my own. However, i do not agree in treating all the layers equally as some layers in my view had a bigger importance than others. I describe the layers transport and presentation as supporting. ChatGPT also included more details on process controls.
I agree with most of the mitigations and defense-in-depth structure that ChatGPT generated, though my report focused more on practical measures.

6.0.2 Use of ChatGPT in the research

ChatGPT was used as an assistant to structure paragraphs rephrase sentences, and propose set up. It was particularly useful to compare explanations against official reports and structuring the report making the drafting faster. It could not replace verification with documented sources, but with clarified and refined prompts it made the answers more aligned with the requirements.

Bibliography

[1] America's Cyber Defense Agency. Russian state-sponsored and criminal cyber threats to critical infrastructure, 2022.

[2] America's Cyber Defense Agency. Tactics, techniques, and procedures of indicted state-sponsored russian cyber actors targeting the energy sector, 2022.

[3] MIRE ATTCK. Triton safety instrumented system attack, 2024.

[4] MITRE Corporation. Atlas matrix, 2025.

[5] Cybersecurity and Infrastructure Security Agency. Mar-17-352-01 hatman—safety system targeted malware (update b). Technical report, CISA, 2019.

[6] Federal bureau of investigation Department of justice. Triton malware remains threat to global critical infrastructure industrial control systems (ics), 2022.

[7] Martin Giles. Triton is the world's most murderous malware, and it's spreading, 2019.

[8] Dragos Threat Intelligence. Temp.vales, 2019.

[9] Alexandre Mundo. Triton malware spearheads latest attacks on industrial systems, 2020.

[10] NCSC. Triton malware targeting safety controllers. Technical report, National Cyber Security Centre, 2017.

[11] Hope Trampski. The triton malware attack. 2024.