TRITON/TRISIS attack

Julia Vister

September 2025

Contents

1 Introduction to TRITON/TRISIS **2**

2 Evidence and Incidents **4**
- 2.0.1 Official and Systematic Reports 4
- 2.0.2 Documented & Industrial Incidents 5

3 Technical Analysis **6**
- 3.0.1 Cyber Kill Chain 6
- 3.0.2 OSI 7-Layer Model 7

4 Mitigation Steps: Defense in Depth **9**
- 4.0.1 Reconnaissance 9
- 4.0.2 Weaponization 9
- 4.0.3 Delivery 10
- 4.0.4 Exploitation 10
- 4.0.5 Installation 10
- 4.0.6 Command and Control 10
- 4.0.7 Actions on Objectives 11

5 Would it be the same today? **12**

6 ChatGPT **13**
- 6.0.1 Question 3 and 4 using ChatGPT 13
- 6.0.2 Use of ChatGPT in the research 13

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, specifically Saudi Arabia, 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 at a plant. SIS does an immediate shutdown on a critical system, or parts of it, if a dangerous condition is detected, and therefore protects human life. [11] Triton is not a standalone piece of malware, but a a custom built attack framework specifically made to target Triconex safety controllers.[4] It 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. [6] It is believed that the malware reads/writes programs, functions and the query state of a SIS controller. [7] [12]

There has not been a lot of progress in investigating the attack, other than diagnosing the attack when the it happened 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 as 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. [12] It is also believed that they gained access to the network through spear phishing, but no evidence is found of this. [11] 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 (TsNIIKhM). [14] Triton was therefore attributed to TsNIIKhM as a Russian government-controlled research institution under Rus-

sia's Ministry of Defense (MOD).This institution supports the Russian army with cyber capabilities and research. [7][7] They are also known as Temp.Veles and XENOTIME. [9] [3] All reports treat Triton as a single, custom ICS attack framework, and there are no publicly known variants or deployments beyond the original.

The attackers apparent goal was to gain remote control of the plant's safety system and to manipulate or disable the SIS, allowing them to interfere with normal protective functions.[14] FireEye/Mandiant's analysis framed the attacker's options against an SIS in three categories [4]:

- Force a false positive trip to shut down a process to cause operational disruption and financial loss.
- Reprogram the SIS to allow unsafe conditions to persist, increasing the risk of equipment, environmental or human harm.
- Combine SIS compromise with direct manipulation of the DCS to create and sustain an unsafe state or hazard, impacting human safety.

The attackers invested a significant time in both developing and testing the Triton's controller logic against the target environment, repeatedly attempting to deliver functioning control logic. These attempts failed in part because of a conditional check in one of the attack scripts and another safety system detected modified logic causing the controllers to enter a fail-safe shutdown. [11] As a result, the attack caused a plant shutdown but did not cause any physical damage. If the attackers had succeeded, achieving the goal of shutting down the SIS, the plant's last line of defense would have been gone. Worst case scenario if succeeded is the release of toxic hydrogen sulfide gas (H2S) which could result in a massive explosion as H2S is extremely combustible. This again would lead to filling the air in the surrounding areas with toxic chemicals. [14]

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. [12]
- **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. [6]
- **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. [7]
- **MITRE ATTACK (2024):** Documents Triton's attack lifecycle and associates it with the threat group TEMP.Veles.[3][9]
- **FireEye/Mandiant (2017-2019):** First discovered the malware and published several reports attributing it, while also providing detailed analysis of the framework's architecture, tools and TTPs. [13] [4] [10]
- **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. [8]
- **Trellix (2020):** Contributed to analysis of how TRITON fits into a broader aspect of ICS-targeting malware and discussed its sophistication and persistence. [11]
- **Purdue University (2024):** A more recent academic perspective on the attack. Summarizes the attack, attribution, and consequences in the context of critical infrastructure protection. [14]

2.0.2 Documented & Industrial Incidents

To date, the Triton malware has been publicly linked to a single confirmed incident: the attack on the petrochemical facility in Saudi Arabia in August 2017. This incident represents both the documented and industrial incident associated with the malware.
The main parts of the incident consisted of six Triconex controllers and the Triconex Safety System. The logs showed that unauthorized configuration changes was pushed to the controllers. These types of changes can only be made by someone physically at the controllers as you have to insert a key into it and turn it into PROGRAM mode. When the controller is in PROGRAM, someone in the control room can push the configuration change. Then, turn the key back to RUN mode. The computer in the control room had an unauthorized RDP which gave the attackers remote access and most likely a worker had forgotten to switch the mode back to standby. The configuration changes caused a fail safe state on the SIS controllers, which triggered a plant shutdown. When incident responders arrived at the plant, they found the compromised computer where the binaries library.zip (the payload package) was found as well as the main executable leveraging it, trilog.exe. [4] An earlier, smaller incident two months prior to this had affected only one controller, which most likely indicated they were testing their framework before a full deployment. Later, after forensic investigators found the compromised engineering workstation containing the Triton files, the US and UK government advisories attributed 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]

Chapter 3

Technical Analysis

3.0.1 Cyber Kill Chain

TRITON follows a classic persistent threat kill chain that is notable for its long reconnaissance and testing to gather information about the system. The gathered knowledge includes: Triconex controller models and firmware, engineering documentation, and the network topology connecting the workstations to the SIS. This enabled an offline weaponization phase that was 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 is believed to be through a entry vectors, for example spear-phishing and stolen credentials. From this the attackers 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, and exploited engineering tools, protocols and credentials to gain the privileges required to reach the SIS controllers.[3] During the installation of the malware, the attackers staged Triton components in the workstation and deployed controller payloads,trilog.exe and library.zip. [4] The attacker maintained persistence on the workstation and established command-and-control pathways across the IT-OT boundary by using remote tools like WMImplant, RDP jump boxes and backdoors to coordinate timed operations. [3] The last stage, actions on objectives, the attackers used TriStation sessions to reprogram SIS logic and send unauthorized command messages to the Triconex controllers, resulting in a triggered fail-safe condition that shut down the plant. [3] See figure 3.1.

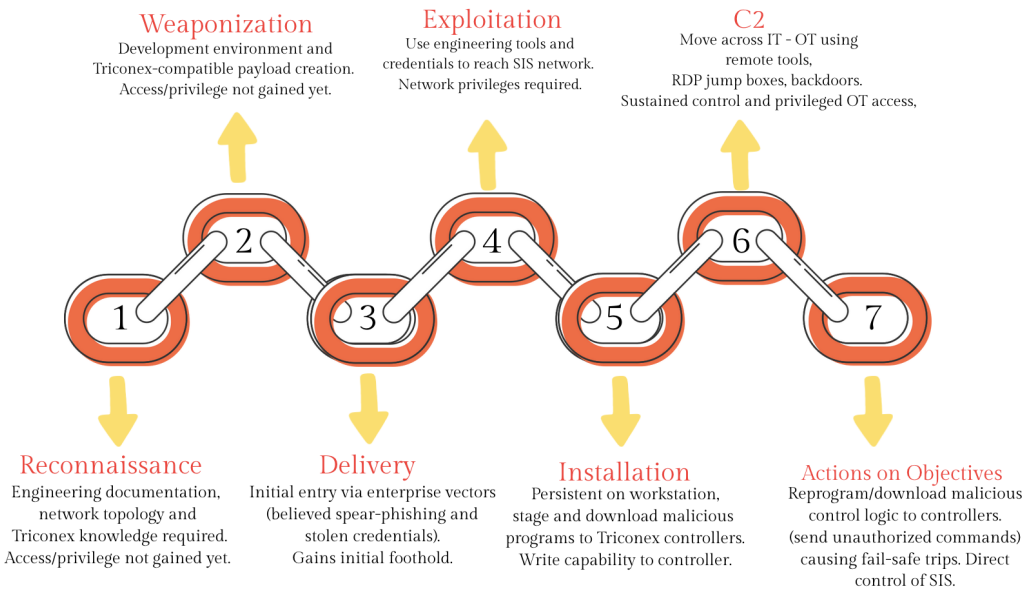

Figure 3.1: Cyber Kill Chain of TRITON

3.0.2 OSI 7-Layer Model

TRITON activity is documented from Physical layer and upwards. Reports provide evidence of how the malware operated across Layers 2-7, with the Data Link, Network, Session and Application layers being the main focus of reconnaissance, movement, and payload delivery.

- **Layer 1 — Physical**
  Triconex safety controllers require a physical key switch (RUN/PROGRAM). Controller 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 Protocol) 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. [6] 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

Defense in depth is based on applying multiple, complementary security controls across people, process and technology so that if one layer fails, others still protect the system. The following mitigation steps proposed are mapped to the Cyber Kill Chain stages identified in Section 3, and made from lessons learned from the Triton attack. [4]

4.0.1 Reconnaissance

- People: Train staff not to share plant details on public platforms. Reinforce awareness that even harmless information can aid adversaries.
- Process: Enforce strict access controls, and classify engineering documents and SIS configurations as sensitive. Restrict access to "need-to-know".
- Technology: Implement IDS (Intrusion Detection Systems) to monitor for unauthorized scanning or asset discovery.

4.0.2 Weaponization

- People: Staff is updated on adversary techniques and ICS-focused malware through training.
- Process: Maintain a threat intelligence program to track actor TTPs and update defensive measures accordingly.
- Technology: Detect known tool malware signatures and weaponization patterns by deploying sandbox analysis tools. Maintain up-to-date behavioral detection baselines on workstations.

4.0.3 Delivery

- People: Continuous phishing-awareness training for all employees.
- Process: Security awareness campaigns and periodic phishing simulations.
- Technology: Email filtering and reputation-based blocking, MFA for all external remote logins. Restrict use of removable media between IT and OT.

4.0.4 Exploitation

- People: Train engineers/operators to verify SIS changes or alerts.
- Process: Dual-control approval for any SIS configuration change or controller logic update. Record and monitored all sessions.
- Technology: Segregate safety system networks from process control and information system networks. Ensure workstations are not dual-homed to DCS and IT networks. Use unidirectional gateways for systems that require data from SIS. Strict access control and application whitelisting on any workstation that reach SIS over TCP/IP. Monitor ICS network traffic.

4.0.5 Installation

- People: Train operators to recognize anomalies such as unexpected trips, controller reboots, or modified logic.
- Process: Frequent integrity checks of controller logic, firmware, and configuration baselines. Implement change management procedure for SIS key-switch position changes.
- Technology: Endpoint detection and response solutions on workstations. Use hash-based verification for controller logic to detect changes.

4.0.6 Command and Control

- People: Remote access logs reviewed and investigation of alerts by administrators.
- Process: Remote session logging, and and enforce approval workflows for remote maintenance.
- Technology: Strong network segmentation between IT and OT, block direct RDP connection to SIS network, and monitor command channels for anomalies using IDS.

4.0.7 Actions on Objectives

- People: Engineers and operators trained to verify and respond to abnormal SIS behavior, such as unexpected shutdowns and logic changes.
- Process: Regular incident response drills to improve readiness.
- Technology: Triconex controller keyswitches in RUN mode when not in use, configure alarms for changes in PROGRAM mode, and automated alerts for logic modifications.

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 possesses the patience and resources to remain stealthy for years, meaning no single defensive layer is enough. Continuous training, enforced security processes, and the integration of monitoring technologies across IT and OT layers provide the most effective defense-in-depth strategy.

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. [5]

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.[5]

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] Marina Krotofil Dan Scali Nathan Brubaker Christopher Glyer Blake Johnson, Dan Caban. Attackers deploy new ics attack framework "triton" and cause operational disruption to critical infrastructure.

[5] MITRE Corporation. Atlas matrix, 2025.

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

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

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

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

[10] FireEye Intelligence. Triton attribution: Russian government-owned lab most likely built custom intrusion tools for triton attackers.

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

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

[13] Daniel Kapellmann Zafra Dan Caban Steve Miller, Nathan Brubaker. Triton actor ttp profile, custom attack tools, detections, and attck mapping, 2019.

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