Mitigating software supply chain attacks: tracking and containing data manipulation attacks

(Allison Barnard Feeney abf@nist.gov, Sylvere Krima sylvere.krima@engisis.com, krimas@nist.gov)

The software supply chain: a threat to trustworthy data and decisions

Software can be summarized as a set of instructions that consumes and produces data from/to other software (or people), creating a data supply chain. Any disruption of the software supply chain itself will also disrupt that data supply chain and potentially compromise the data confidentiality, integrity, and/or availability. This disruption at the data-level has become a significant challenge with the recent and fast-paced digital transformation of private industries and public organizations.

The digitalization of industries and public organizations has led these entities to become data-centric [DoD 2021] and reliant on data and analytics to drive business decisions. As data became the primary fuel to economic growth and business strategies, it also became a significant source of concern. Data confidentiality, integrity, and availability are security concerns that organizations face daily and must address to support their data-centric nature. With most efforts focusing on protecting data confidentiality (e.g., against data theft) and supporting its availability (e.g., against ransomware), data integrity is trailing behind despite its importance [Accenture 2019, Booz Allen 2020] and it being a priority to most organizations [KPMG 2018, Deloitte 2019]. Data integrity is key to delivering trustworthy data and analytics that can be safely consumed by organizations. Data integrity guarantees that malicious actors have not compromised the data (e.g., through manipulation attacks) [MITRE 2020], thus disrupting decision-making processes that rely on it. This type of disruption is also a significant concern to the nation's economic and political systems [DoD/DHS 2016, Air Force 2017].

Embracing a data-centric approach relies on maximizing data sharing across IT systems and organizational silos [DoD 2021], supported by a large number of software and data flows (see **Figure *1***). These changes are opportunities for malicious actors for whom: 1) each software is a new potential way to manipulate data, 2) each data flow is a vector to propagate corruption across systems and organizations, 3) each corrupted data increases the chances of misleading decision makers and compromising organizational decision-making processes [DoD 2014]. These opportunities grow exponentially as organizations transition from monolithic software architectures to microservice architectures, drastically increasing the number of software and data flows, outstripping the capacity to track them manually.

Trusting your data: tracking the software data supply chain.

In order to make trustworthy decisions [KPMG 2018], organizations need solutions to prevent and quickly contain data corruptions once they have been reported. Our proposal relies on developing tools and methods for software providers and users to document and track their IT systems and how data flows between them (see **Figure *1***). Each software provider would deliver software data exchange specifications (i.e., how data maps from one system to another [Krima 2020]) in standard form that could be used by the end users, after an incident has been reported, to: 1) identify where the corrupted data has already propagated to (purple in **Figure *1***) and potential future propagations, 2) quarantine compromised software and users to contain the propagation, and 3) identify and track compromised business decisions. These specifications can also be used by end users to identify systems that would become super spreaders once corrupted and take the appropriate preventive measures.

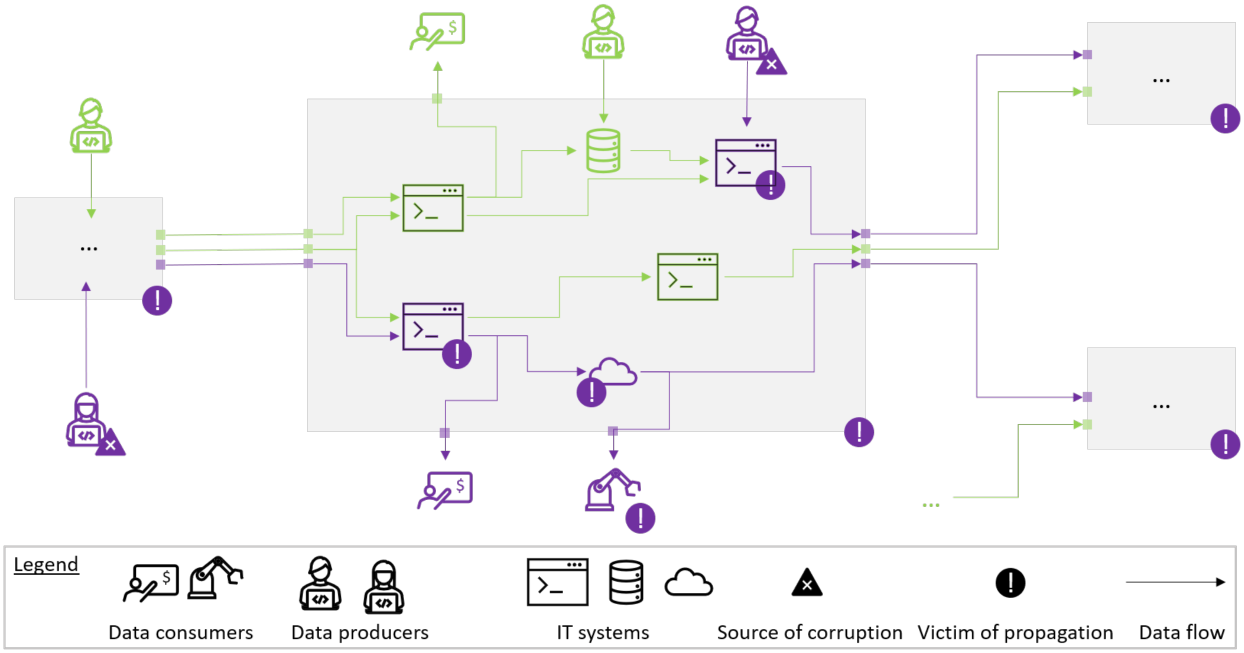

Figure 1. Graphical illustration of data corruption and its propagation across organizations (gray blocks), systems, and users. Purple elements have been corrupted.

These solutions must be open, standardized, and computer interpretable. When it comes to security, a supply chain is only as secure as its weakest element. An open and standardized solution will lower adoption barriers, especially for smaller organizations, and guarantee interoperability between actors of the supply chain, reducing deployment costs. With an average of 280 days to detect and contain a data incident (and 315 days when it was caused by a malicious attack) [IBM 2021], fast detection and containment measures are key, and require computer interpretable solutions that can be automated. Moving under the 200 days threshold is expected to save an average of $1.12 million per incident [IBM 2021].

Why NIST

Standards and guidelines to document and track data flows across software and organizations, in support of the **Executive Order section 4 (e)**, is not a new challenge to NIST. Over the past few years, NIST researchers (from the Engineering Laboratory - System Integration Division – PoC: abf@nist.gov) have been developing methods and tools to address some of these challenges [Hedberg 2020, Krima 2019, Krima 2020].

NIST experts understand the complexity and requirements associated with developing open and standardized solutions. Such solutions lower adoption barriers and are key to achieving the NIST mission to improve and promote US competitiveness, by putting cutting-edge cyber security methods and tools in the hands of US entities.

Finally, NIST technical work is a neutral foundation, free of commercial influence, that can support law enforcement agencies and law and policy makers in potential efforts to: 1) redefine and align software acquisition federal requirements and procedures with cyber security concerns, 2) design federal response strategies to software supply chain attacks and similar threats, and 3) track and contain data manipulation across the federal IT infrastructure.

References

[Accenture 2019] "The cost of cybercrime", Accenture, 2019. Accenture-2019-Cost-of-Cybercrime-Study-Final.pdf

[Air Force 2017] "US Policy Response to Cyber Attack on SCADA Systems Supporting Critical National Infrastructure", Scott A. Weed, Air Force Research Institute, May 2017

[Booz Allen 2020] "Data manipulation risk", Booz Allen Hamilton, 2020. Life Sciences and Clinical Technology: Seven Trends for 2019 (boozallen.com)

[Deloitte 2019] "The future of cyber survey 2019", Deloitte, 2019. The future of cyber survey 2019 (deloitte.com)

[DoD 2014] "Information Operations", Joint Publication 3-13, Joint Chiefs of Staff, Department of Defense, 20 November 2014

[DoD/DHS 2016] "Critical Partnerships: DHS, DoD, and the National Response to Significant Cyber Incidents", Dr. Andy Ozment (Department of Homeland Security), Tom Atkin (Department of Defense), 2016. DOD-DHS-Cyber_Article-2016-09-23-CLEAN.pdf (defense.gov)

[DoD 2021] "Memorandum For Senior Pentagon Leadership Commanders Of The Combat Ant Commands Defense Agency And DoD Field Activity Directors – Creating Data Advantage", Deputy Secretary of Defense, Department of Defense, 05 May 2021. Deputy Secretary of Defense Memorandum

[Hedberg 2020] "Recommendations on Ensuring Traceability and Trustworthiness of Manufacturing-Related Data" – Hedberg T., Helu M., Krima S., Barnard Feeney A. – NIST Advanced Manufacturing Series 300-10, July 2020, https://doi.org/10.6028/NIST.AMS.300-10

[IBM 2021] "Cost of a Data Breach Report 2020", IBM Security, 2021.

[Krima 2019] "Securing the Digital Threat for Smart Manufacturing: A Reference Model for Blockchain-based Product Data Traceability" – Krima S., Hedberg T., Barnard Feeney A – NIST Advanced Manufacturing Series 300-6, February 2019, https://doi.org/10.6028/NIST.AMS.300-6

[Krima 2020] "Toward Model-Based Integration Specifications to Secure the Extended Enterprise" – Krima S., Toussaint M., Barnard Feeney A. – ASTM Smart and Sustainable Manufacturing Systems, Vol. 4, No. 1, 2020, https://doi.org/10.1520/SSMS20200022

[KPMG 2018] "Guardians of trust", KPMG, February 2018. Guardians of trust _FINAL WEB.pdf (assets.kpmg)

[MITRE 2020] "Data manipulation", MITRE ATT&CK framework, 02 March 2020. Data Manipulation, Technique T1565 - Enterprise | MITRE ATT&CK®