STORAGE DEVELOPER CONFERENCE
SDC 22
Fremont, CA
September 12-15, 2022
BY Developers FOR Developers
A SNIA Event

Data Loss Mitigation through 2-Factor Authentication

Presented by

Peter Scott, Senior Principal Engineer, Peter.Scott@ThalesGroup.com

Rajesh Gupta, Director of Engineering, Rajesh.Gupta@ThalesGroup.com

THALES

Overview of Ransomware Attacks

The End Goal of a More Secure Solution

Attacks in Todays' World

- Information needed on what is relevant in todays world of attacks
  - How current solutions fall short
  - Global market value of a solution even if not 100%
  - How attack vectors are introduced to a system
- According to the Federal Bureau of Investigation (FBI), as many as 2,048 ransomware complaints were registered in 2021 alone.
- Malicious emails are up 600 percent due to COVID-19. (ABC News, 2021)
- The average downtime a company experiences after a ransomware attack is 22 days. (Statista, 2021)

Attacks in Todays' World

- Traditional access control on confidential and sensitive files is not enough
  - Unfortunately, Credentials can be stolen, shared, bought or hacked.
  - Once rough applications (malwares) gain entry, the threat actors will often look to compromise privileged access credentials to further infiltrate your network and steal sensitive data.
- Signature based verifications falls short
  - Zero day attacks are becoming more sophisticated
  - In most scenarios, administrators do not even know for weeks that their network or data is under attack

What we are not trying to solve

- Signature based recognition
  - This is a moving target and requires a huge team to constantly recognize and update ransomware signatures
  - Constantly pushing new sets of signatures to every client
- 100% Data Integrity
  - Requires secondary caching of content; i.e. CoW implementation, etc.
  - Overly complex solutions for an already complicated layer of security
- Full system security
  - Protection on a per directory basis so only ‘important’ data is protected. Though can protect against Petya-like attacks of MBR and Partition Tables

What we are trying to solve

- Zero-Day Attacks
  - Through a combination of a ‘learn mode’ and access heuristics, build a solution that can recognize rogue processes accessing specific data sets
- Almost 100% Data Integrity
  - Denying access to rogue processes after a short time to minimize data loss
  - ‘Some’ files may be lost but the goal is to minimize this delta
- Protect only specific data sets
  - By focusing on specific directories containing the protected data, we can narrow down the access patterns and data usage to a specific set of heuristics

Using Multi-Factor Authentication

A High Level Look into the Solution

Multi-Factor Authentication

- Allows for a method to validate a users identity … in most cases
  - Standard 2FA, when built around a robust authentication backend, minimizes the chances of a rogue user getting through
  - Used by a wide variety of services on the web such as bank access, medical records, etc.
  - There are potential system-in-the-middle attacks on 2FA but they have yet to be widely exploited

How to Leverage MFA in a Robust Solution

- Recognition
  - At the core of any solution is recognizing when a process is acting out of character
  - This requires some level of understanding of what is in character
- Collecting Heuristics
  - As part of the recognition phase, User accessing the data and processes running on the system will be scored based on their interaction with the protected data sets
  - Some of these scores include, but not limited to:
    - The type of data, encrypted vs non-encrypted, being read and written
    - Metadata operations on the files
    - User access pattern

How to Leverage MFA in a Robust Solution .. Cont'd

- Trigger Point
  - At some point of the information gathering phase, a trigger point will be hit
  - Any further access on the protected data set will be denied to the user
- MFA … The Validation Point
  - At this point, force the caller to go through 2FA to continue accessing the protected data set
  - Without clear authentication of the user, all further accesses will be denied for the process

How to Leverage MFA in a Robust Solution .. Cont'd

- Fallback and Exempt List Interface
  - Of course there will be specific processes which will act out of character on a given data set
    - Backup and restore applications
    - Xcopy-like applications
  - Business decisions can be implemented through policy to allow specific processes from being denied access to the data set
  - Additional checks can be used to ensure these exempt applications have not been compromised such as hash calculations on the binary which are checked at process load time

A Detailed Look into the Implementation

Achieving Our Solution in a Windows Environment

- Base architecture is achieved through a Filter Manager style File System Filter driver
  - Callbacks added for open/create, metadata and I/O operations as well as cleanup and close processing
  - A process create and image load callback is setup
    - This allows us to follow the actions of each individual process

Our Solution .. Cont'd

- Learn Mode Infrastructure
  - There will be a period of time post installation of the product where heuristics are gathered for each data set on the system
  - The information gathered is per data set and not tied to a specific process running within the data set
    - This allows for a more generic application of the heuristics across a given data set for any process running
  - Some of the information collected include:
    - Open and create operations per second
    - Overall IO throughput as well as IOs per second of the dataset
    - Metadata operations such as delete, rename, file size changes, etc
    - Average data entropy read and written per fixed size block

Our Solution .. Cont'd

- Learn Mode Infrastructure … continued
  - The collected information will be stored in the registry for later retrieval
    - Administrators can review this information at any point for further analysis
  - Specific process names, while gathered and referenced in the statistics, it is not used for the application of the information, it can be used by administrators to possibly exempt specific processes from the actions of the product
  - If a process is exempted by an administrator, any data collected associated to this process will not be used for the general application of the data for protection of the data sets
  - Once the collected data set is approved by the administrator, calculations will be performed on the data to ease the application of the heuristics during run time actions

Our Solution .. Cont'd

- Run-time Protection
  - At the time of process launch, the driver will determine if it has been marked as exempt, or not
    - If an administrator has marked a specific process, by name, as exempt from run-time processing, at the time of process load the hash will be generated and compared against a previously known value. This will ensure an exempt process is not modified outside of the run-time protection framework
  - For a process not in the exempt list, control structures will be allocated for collecting run-time information on this process
    - Weighted averages for statistics collected during the learn mode processing will similarly be collected for each individual process
    - This information will be monitored by an external thread

Our Solution .. Cont'd

- Run-time Protection … continued
  - Process specific information will collected in-line to the operations of the process but the actual calculations and comparisons will be performed in an external thread to ease the performance burden of this processing
  - Based on these comparisons to the collected heuristics, a score will be applied to each process
    - The individual data point scores can add or subtract from the overall score of the process
  - If a preset threshold is exceeded for a process, secondary actions will be taken to potentially exclude the process from accessing more data within the dataset
    - These thresholds are determined after the learn mode data has been analyzed.
    - An administrator has the option to set thresholds for specific processes, if needed. In general, all processes as assigned the same threshold

Our Solution .. Cont'd

- Run-time Protection … continued
  - At the time of the trigger point for a given process, an administrator can configure the action to be taken
    - The default action will be to deny any further access by the process within the protected data set
      - This would be enforced by the file system filter driver from performing any file system operations
      - A request to have the caller perform a form of 2FA to ensure the caller is allowed to continue accessing the data
    - An administrator can configure this action to be notified via email, text, etc. vs a proactive action as above
  - Further actions by an administrator can be taken for processes which have crossed the threshold
    - They can add the process to the exempt list
    - Adjust the threshold for the process

Conclusion : Mitigate the impact of Ransomware

- Reduce the radius
  - Reduce your blast radius by limiting access to critical data so that only those who require access have it.
- Implement a ‘zero-trust’ security model
  - Assume your perimeter defenses will fail and make sure everything within is still safe and secure.
  - Implement Zero Trust Security Model and authenticate all users and devices that connect to your network every time they connect
  - Authenticate all users before they access critical data

Please take a moment to rate this session.

Your feedback is important to us.

THALES

STORAGE DEVELOPER CONFERENCE SDC 22