napatech

"ARTICLE"

CYBER SECURITY AND PACKET CAPTURE: MAKING THE CONNECTION

Cyber Security and Packet Capture: Making the Connection

By: *Peter Ekner, Chief Technical Officer, Napatech*

Yahoo recently suffered a major data breach in which a copy of undisclosed information on 500 million user accounts was stolen. Though detailed information on exactly what happen is yet to be revealed, previous data breach incidents tell us that the investigation will be complex and time-consuming.

What makes cyberattacks so tricky is that there isn't a traditional, physical crime scene where evidence can be collected for investigation. Instead, we are facing a crime scene built from a complex structure of servers, networks and applications, scattered across many different geographical locations.

Another hindrance to cyber security forensics is the fact that servers, networks and applications only provide a partial and reduced set of evidence. For example, a log file from a server can show the health of the server and applications running at any given time, but it will not be able to tell exactly what information was exchanged with other servers, networks or applications. Similar arguments can be made for log files originating from networks and applications.

To increase the amount of evidence, we need to shift our focus away from these devices and onto the actual information traversing our networks. By collecting this type of information, we can reconstruct a complete picture of what occurred by deploying full packet capture capabilities at strategic points across the network infrastructure.

Ensuring Zero Packet Loss

The first critical challenge to address regarding reliable packet capture for forensic evidence is to ensure every single packet is captured. Imagine discovering during a forensic investigation that you are missing the single piece that could complete the puzzle.

A high-speed, uncompromised packet capture solution is needed to be able to capture every single packet, no matter the packet size and packet patterns, at the maximum network operating rate. For instance, when doing packet capture on a fully utilized 10Gbps network link, 1.23Gbyte of new packet data must be written and up to 14.88 million new records must be added to the packet capture database every second.

Rapid Retrieval of Relevant Information

The second critical challenge is finding the relevant information for our forensic investigation in a packet capture database with a size of several hundred TBytes or even PBytes and with billions of individual records. This can best be described as finding the proverbial needle in the haystack.

However, it is infeasible to go record by record through the entire packet capture database. Instead, the packet data must be indexed as it is written to the packet capture database to enable fast searching. The most common ways of indexing the packet data is on reception time, addresses, protocol number and port numbers. Indexing by reception time will enable us to quickly find all packet data captured within a certain timeframe; indexing by addresses, protocol number and port numbers will enable us to quickly find all packet data exchanged by either one user or between two users. The various types of indexes can also be combined, allowing us to search fast for all data exchanged between two parties within a given time window.

Indexing packet data on reception time, address, protocol number and port numbers is an efficient way to find packet data for a forensic investigation quickly. Efficiency can be improved further by

associating every packet data origination from a given communication session with a unique session ID and indexing all packet data by their unique session ID. Doing this will quickly find all data packets belonging to a given session between two entities, such as in a specific YouTube video playback.

In addition to overcoming the challenge of finding data packets of interest quickly, it is also important to get the relevant packet data retrieved from the packet capture solution and into the hands of the forensic network security team for analysis as fast as possible.

Fast retrieval is important for several reasons. Imagine investigating a possible security breach and quickly identifying some suspicious packet data in the packet capture database, only to then spend several hours retrieving the suspicious packet data. Firstly, this will prevent the forensic network security team from making progress until the retrieval process is complete. Secondly, there is a chance that the suspicious packet data will be overwritten by newly captured packets before the retrieval process is done.

Gathering Forensic Evidence

Having packet capture capabilities is like having a time machine. We now have a complete picture of what happened 10 minutes, one hour, one day, one week, one month or one year ago on the network. The big question is: How far back in time must we be able to travel?

No CIO wants to be in a situation where a possible cyber security attack cannot be investigated due to insufficient packet capture history. The recent Target breach showed us that the attackers were present in the company's IT infrastructure for more than 200 days before the data breach was discovered and a detailed investigation was initiated.

An organization needs to determine how much data storage capacity it needs. It does this by obtaining different levels of packet capture history and determining how much packet data an average organization or enterprise is generating. Let us assume the small/medium enterprise will generate an average network load of 750Mbps across a 24-hour window and a large enterprise will generate 5Gbps under the same conditions.

Using that information, we can calculate the minimum required data storage capacity for one day, one week, one month and one year of packet capture storage.

[image: packet capture storage size]

Because most of today's packet capture solutions can scale to 1000TB of data storage, that equals roughly four months of packet data history for a small/medium enterprise and less than a month for a large enterprise.

To expand the packet data history even further, the packet data can be compressed before it is written to the packet capture database. The effect of compressing the packet data depends on the selected compression algorithm and the content of the data packet, as certain data is more suited for compression than other. Standard network packet data have a typical compression ratio of three; hence, compression can triple the packet capture history.

Counting the Cost

In the ever-increasing battle against critical data breaches, deploying packet capture capabilities is an extremely effective weapon. But as described above, a few challenges must be addressed in order to have a successful packet capture solution.

It’s also important to consider the financial investment required. The two biggest cost drivers are the number of packet capture systems and the size of packet data history. A single packet capture system can cost up to US $250K, depending on the exact system configuration, and a storage solution in the PByte range could cost more than US $1M.

Buying packet capture capabilities is akin to buying life insurance. You want to have the best possible coverage given the amount of money you are ready to spend. Surely the Target executive management team would have preferred such a cyber security “life insurance” defense solution that could have given them a quick forensic investigation of the reported data breach. The rest of us can learn from that story and avail ourselves of every possible cyber security advantage.

About the author:
Peter Ekner is CTO at Napatech and manages the Copenhagen team. He has more than 15 years of experience working with the design of products and components for network equipment.