Vulnerability Management Solutions are Flawed, Leaving your Enterprise at High Risk

The lack of advanced network endpoint correlation in your assessment technology introduces bedlam

December 2014
Gordon MacKay
EVP/Chief Technology Officer
Digital Defense, Inc.

Digital Defense
9000 Tesoro Drive, Suite 100 | San Antonio, Texas 78217 | www.ddifrontline.com

INTRODUCTION

Is the State of Your Security an Illusion?

Many organizations are unknowingly at risk of a devastating security breach.

Why? A false sense of security and misplaced faith in highly touted scanning tools that provide misleading information to the user. Could it be that the reason breaches are being reported almost daily now is due to inaccurate results organizations are relying on to make critical security decisions? Millions of dollars are being spent on technology that does not have the capability to accurately reconcile vulnerability scanning results from scan to scan, opening the door to crippling breaches.

Perplexing Scanning Issues that Can Lead to Breaches:

- A vulnerability scan declares more assets than you really have…
- Scan results show that vulnerabilities have been fixed when you know they have not…
- Scan results show vulnerabilities that don't exist leaving your resources exhausted from chasing ghosts?...

Is your organization chasing issues that do not exist?

Could the real issues be buried under a mountain of inaccurate, misleading information?

Digital Defense, Inc. (DDI) has been conducting vulnerability scans for over a decade. Given this experience, we believe the reason these issues exist is due to inadequate host and vulnerability reconciliation capabilities in tools widely used in the industry today. The term, reconciliation, refers to technology that provides network endpoint correlation (i.e. in this case maintaining accurate records of a host's identity and security posture over time). The lack of proper reconciliation technology generates misleading results and even worse, an inaccurate view of the organization's security posture.

Our Research Speaks for Itself

DDI has research demonstrating how other vendors are missing the mark with tracking methodologies that are based on limited host characteristics. Most vulnerability scanning vendors only track three to five host characteristics. This is not enough for accurate security.

Tracking hosts only by IP address, hostname, or even MAC address is folly at best and negligence at worst. There is too much "churn" in most networks, over 40% in some cases, for such a limited number of variables to be monitored. Users experiencing discrepancies should begin questionings the accuracy of vulnerability scan findings.

- How do you know results are accurate?
- Are you sure the hosts aren't changing? If so, how?
- Why are my host results not matching up?
- Why are hosts showing up more than once when it is the same system?

DDI has been working the past 18 months to gather insight to help answer the tough questions. Through research, data mining and competitive intelligence we've been able to test, validate and verify our scanning technology.

- Did you know that on average 20% of your hosts will change in some fashion? (IP Address, hostname, MAC address, services, etc.)

- Did you know that 20% of mismatches across a three month time space on a network with 1000 hosts could lead to variances on up to 200 hosts. Additionally, assuming that each host has at least two vulnerabilities; that's 400 vulnerabilities that your teams will needlessly research and address.

While other vendors may track three or four host characteristics, DDI monitors and evaluates up to **TWENTY** host characteristics. These additional characteristics help ensure that the scan results used by information security and IT teams are accurate and provide trustworthy historical host perspectives. This is something DDI does that others do not.

The white paper enclosed we believe holds the answers to improved security by revealing the weaknesses seen in many other Vulnerability Management services. Although the Vulnerability Management term (hereafter referred to as VM) may be used to refer to a more encompassing process which includes VM technology, as well as automated and manual business processes surrounding this, VM is used within this whitepaper to refer only to products and services that are currently available on the market today and which perform automated vulnerability assessments (also referred to as scanning).

We take a purists look at vulnerability scanning and management efforts and explore industry standards and the opportunity for improvement by leveraging a nextgen algorithm and advanced reconciliation.

The Issue

Today's Enterprise organizations are being misled with regard to their security risk exposure, and are in serious danger of becoming victims of security breach events. The automated vulnerability management (VM) solutions and products that are central to every Enterprise information security program, and which are essential in gauging network security information risk, contain a serious "hidden" flaw which is now beginning to come to light. This software flaw is interleaved within pattern matching-like algorithms located deep within the foundational core of the most prevalent and widely used automated VM system products and solutions on the market today. As a direct consequence of this flaw, even though these products report a certain level of network security risk, the metric upon which their calculations are based is skewed, resulting in an unintentional gap between the products' intended information risk measurement and the erroneous measurement actually reported. The nature of this flaw causes the resulting gap error to become compounded over time, further undermining the usefulness of the data they provide. Unfortunately, Enterprise security teams are unaware of this problem, and are making decisions and taking actions based upon an inaccurate information security risk gauge. Even more concerning, the numerous vendors providing tools and services containing this fundamental flaw also appear to be oblivious to (or downplaying) the issue and its potentially severe impact on the accuracy and usefulness of the information their products and services provide.

Organizations who find their current solution exhibits this flaw should either procure a replacement solution that employs more robust reconciliation methodology or, alternatively, integrate with a third party solution which can detect and eliminate the errors present in the information provided by their current VM solution.

Should you be concerned about this flaw?

Most of the VM solutions on the market today are ill-equipped to deal with "network churn". A recent study on the prevalence of change to an organization's network asset characteristics, such as an asset's IP Address, Hostname, MAC Addresses and others, and which I share in more detail later in this whitepaper, reveals that these characteristics change far more often than was once assumed. In addition to this, the study also included investigating the various techniques employed by existing VM solutions to attempt to overcome the challenges resulting from such network churn. What we found was astounding. The primitive algorithms within the inner working of the VM solutions supplied by even the largest of the vendors in the space are seriously flawed, and cannot correctly track the findings for the ongoing assessed hosts in the presence of the dynamic network change that our study exposes. As a result of this normal ongoing network churn and the limitations of existing offered VM solutions with regard to tracking asset change, the reality is that the findings portrayed within the "asset views" of the VM systems used by most organizations, including the Fortune 500 Enterprises, are far less accurate than we once believed.

As you will see later in this paper, the weaknesses in these algorithms can lead to seriously incorrect security trend statistics, and even "masking" of important vulnerability information.

You should find this whitepaper of high value if your organization's approach to cybersecurity is aimed at attaining and maintaining a high level of security. The flaw discussed here can seriously hamper your efforts in preventing the compromise of critical systems and information, and severely reduce the value and validity of the information your cybersecurity team gleans from the VM used in your organization. As you will see, the weakness and vulnerability information your team relies upon may be seriously flawed or missing.

If your primary interest in cybersecurity solutions is purely compliance oriented, and where checking boxes for regulatory purposes is the primary goal, this issue will frankly be of little interest to you.

Gauging Information Risk Across Time

Important capabilities required and ideally provided within information security risk solutions and products, including those offered within the network VM space, encompass the ability to gauge the risk associated with an organization's information assets, gauge the change in this risk over time, and ideally, re-evaluate past views of their risk when faced with newly-discovered information applicable to the past.

To place this in real world perspective, let's explore the following use case. Suppose a recently-announced new Zero Day vulnerability affects Apache Web Server releases 2.4.7 to 2.4.9, but not version 2.4.10 and later. You would want to find hosts that may be affected both now (so that remediation efforts can be directed to them) **and in the past** (since these hosts may have already been compromised). In order to determine what assets in your network are impacted, you could perform an automated vulnerability assessment, which typically includes the ability to detect application details such as this. Following the assessment, you could then query your VM system for all hosts that have the problematic versions of Apache Web Server installed. However, if you limit your investigation to this, you will miss some possible candidates which at one point in the past had these affected versions installed, but at present either the application is no longer installed or has been updated to a version that is. Therefore you would also need to query your VM system's past assessments in order to find all of the candidate hosts. The past is therefore a very important part of your overall information risk picture. There are many other use cases to demonstrate this. Since any security program must include an ongoing understanding of risk, it must include recurring assessments over time. The takeaway from this use case is the following:

Vulnerability management (VM) systems must be capable of relating the real world assets that make up your network to the corresponding hosts and devices that were discovered within each independent assessment, both now and in the past.

This concept of relating real world assets to hosts discovered within independent assessments is illustrated in the following figure.

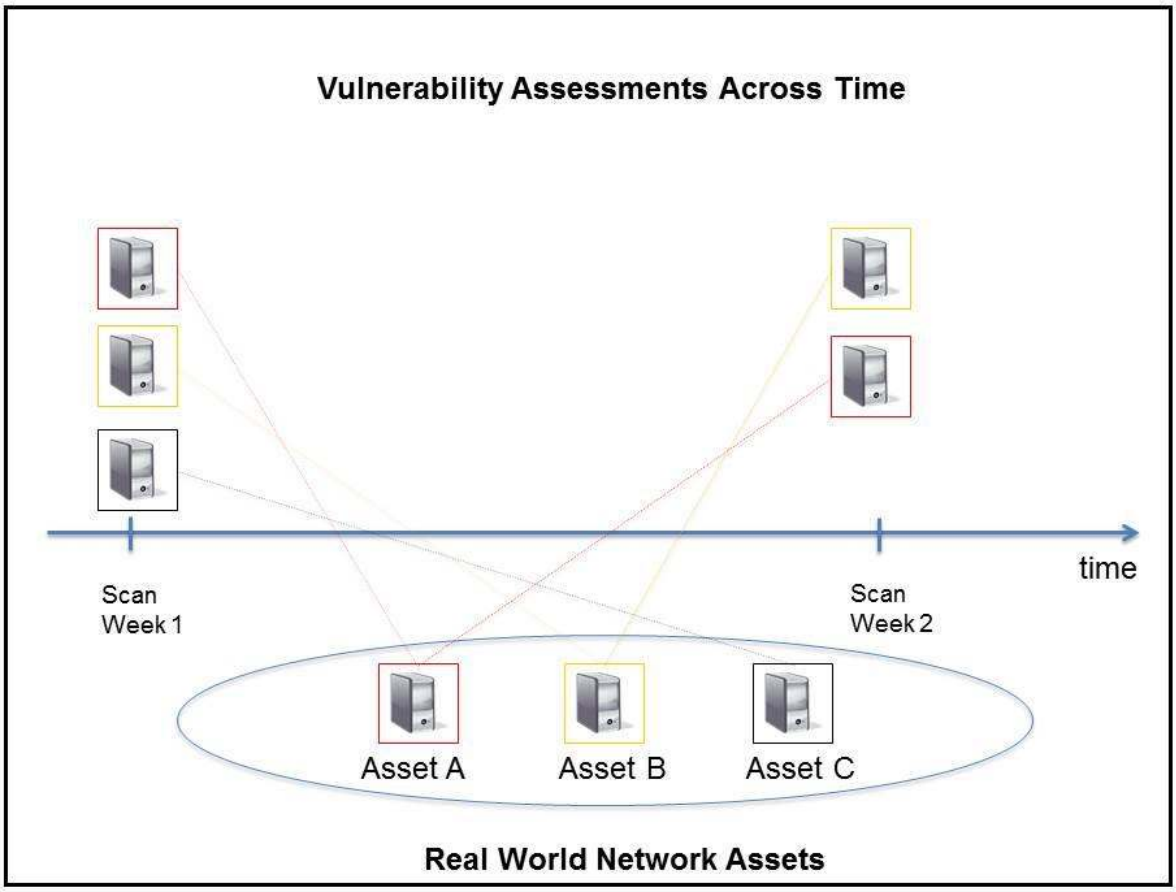


Figure 1 Real World Assets Appearing Over Time Within Different Point-In-Time Assessments

In this example, Assets A and B have been re-deployed or re-configured since week 1, and Asset C has either been removed from the network (either temporarily or permanently) or was powered down/offline at the time of the second test.

Vulnerability Management (VM) Systems Challenged

As I have already explained, most VM solutions on the market today are severely challenged. Accuracy with respect to one "scan" or one vulnerability assessment is one thing, but accuracy with respect to rolled up risk information related to independent assessments across time, which should map directly to risk information related to the real world assets and accurately portray changes in security posture from one scan to the next, is another story. Unfortunately, when VM solutions are compared by industry analysts as well as by prospective buyers, much emphasis is placed on the accuracy of one point-in time sample assessment, but very little emphasis is placed on the accuracy of the historical findings for real world assets, which are in effect the combined rolled up information of the many ongoing point-in-time assessments within any VM process. Yet it is this this very ability to accurately compare results over time that had led the industry to re-name such solutions as *"Vulnerability Management"* rather than *"Vulnerability Assessment"*, since there is a clear need to manage things that change over time. The challenge inherent to VM and which is related to accuracy of findings across time is described next.

Technology Background Related to the Challenge

Before diving into the description of the serious challenge, let's first explore the technology which gives rise to it.

There are different technological options available for discovering assets and the weaknesses resident within the hosts and devices that make up an enterprise network. These include remote network discovery, authenticated discovery, agent-based discovery, and passive inline discovery.

Agent based and authenticated discovery both provide the ability to perfectly track a real world asset to its related discovered information in the various independent assessments conducted for that asset performed across time. Unfortunately, these methods come with high administrative costs. Of these two, agent based discovery entails the highest deployment costs and as such, very few vendors use this method. Although inline discovery is useful and has its advantages, it has limitations as far as the types of information and weaknesses it can discover.

It is for these reasons that the primary and most widely used method is remote network discovery. It is simple to deploy and requires very little to no network IT administration, and its depth of discovery of weaknesses and vulnerabilities is vast. In addition to this primary method, most vendors also offer authenticated discovery method. Since this method involves ongoing IT administration to maintain credentials, and in many cases requires deployment of authentication agents for UNIX and Linux systems, it is used far less frequently than the primary remote network discovery method, and is often only used to assess those parts of the network that are deemed believed to be high risk targets.

> The Challenge: The ability to track the hosts and devices discovered within independent discrete assessments over time to real world assets.

Unfortunately, remote discovery method by its very nature cannot directly "see" what lies within the elements it discovers and assesses. That being the case, solutions which employ it face a complex and misunderstood challenge.

Next I provide a more detailed example to illustrate this challenge and the reasons why its solution is crucial to providing the aforementioned "time related risk awareness" capabilities.

The Challenge - Tracking Assets Across Time

To explain this challenge, I refer to the following figure which illustrates two different and independent vulnerability assessments performed at different points in time, the hosts discovered within each assessment, some of their discovered characteristics, and the assessed hosts' association to their corresponding true real world assets.

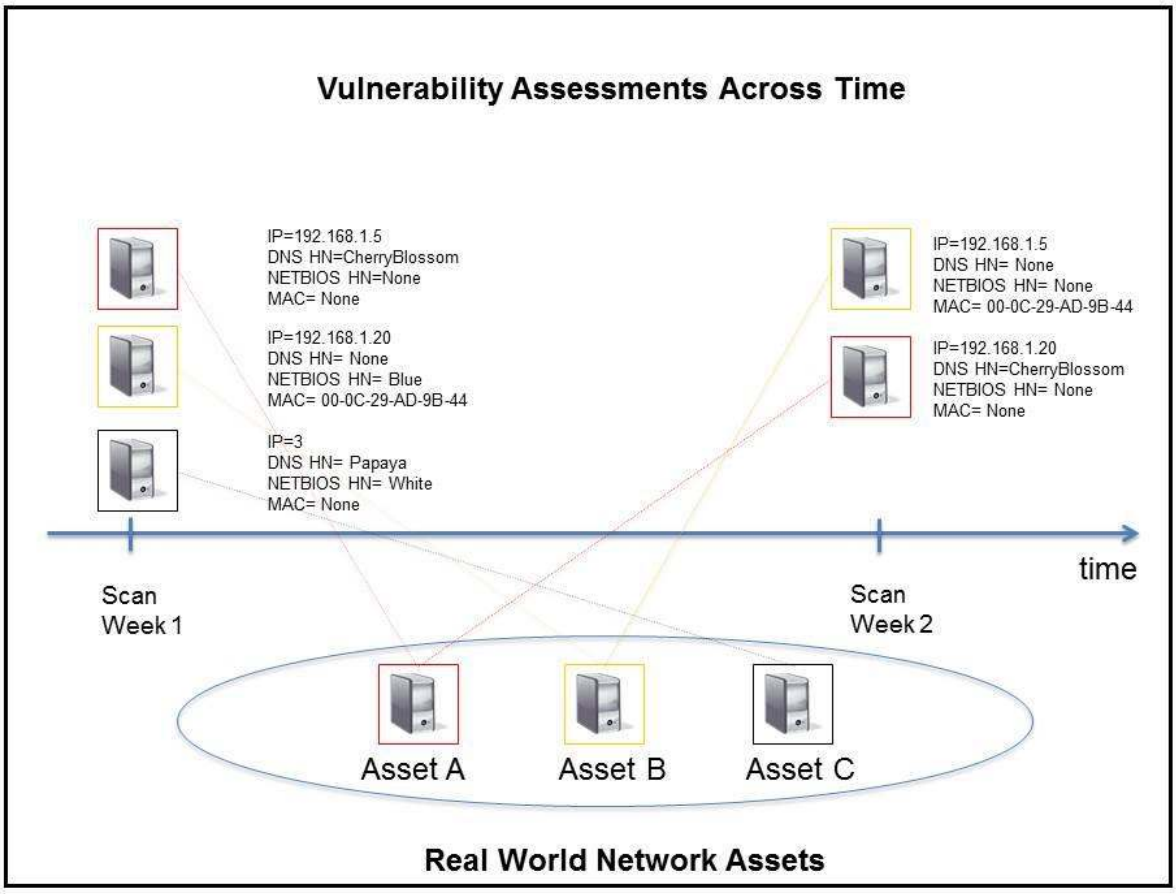


Figure 2 Matching Assessed Hosts to their Correct Real World Assets

The remote network discovery method does not have a presence inside hosts it is assessing, and therefore it must discover the hosts, their characteristics, and their related security weaknesses from an "outside" perspective. This method identifies reaps this information by observing how these assets respond to the Internet messages they receive from a "remote" scanning entity, typically referred to as a "network scanner." Although the remote network discovery method yields enormous amounts of information, it is limited in what it can discover as far as a unique and unchanging piece of information

that it needs to identify in order to track a given asset over its operating lifetime. In essence, a solution that relies on information discovered by way of remote network discovery method is challenged in its attempt to recognize what it "saw" at one point in time, to what it "saw" at a different point in time. There are two primary reasons for this. First, there is no "magic bullet" characteristic present on the entities being discovered which is always discoverable and unique to "latch" onto. Second, those characteristics which are always present and discoverable are subject to change over time.

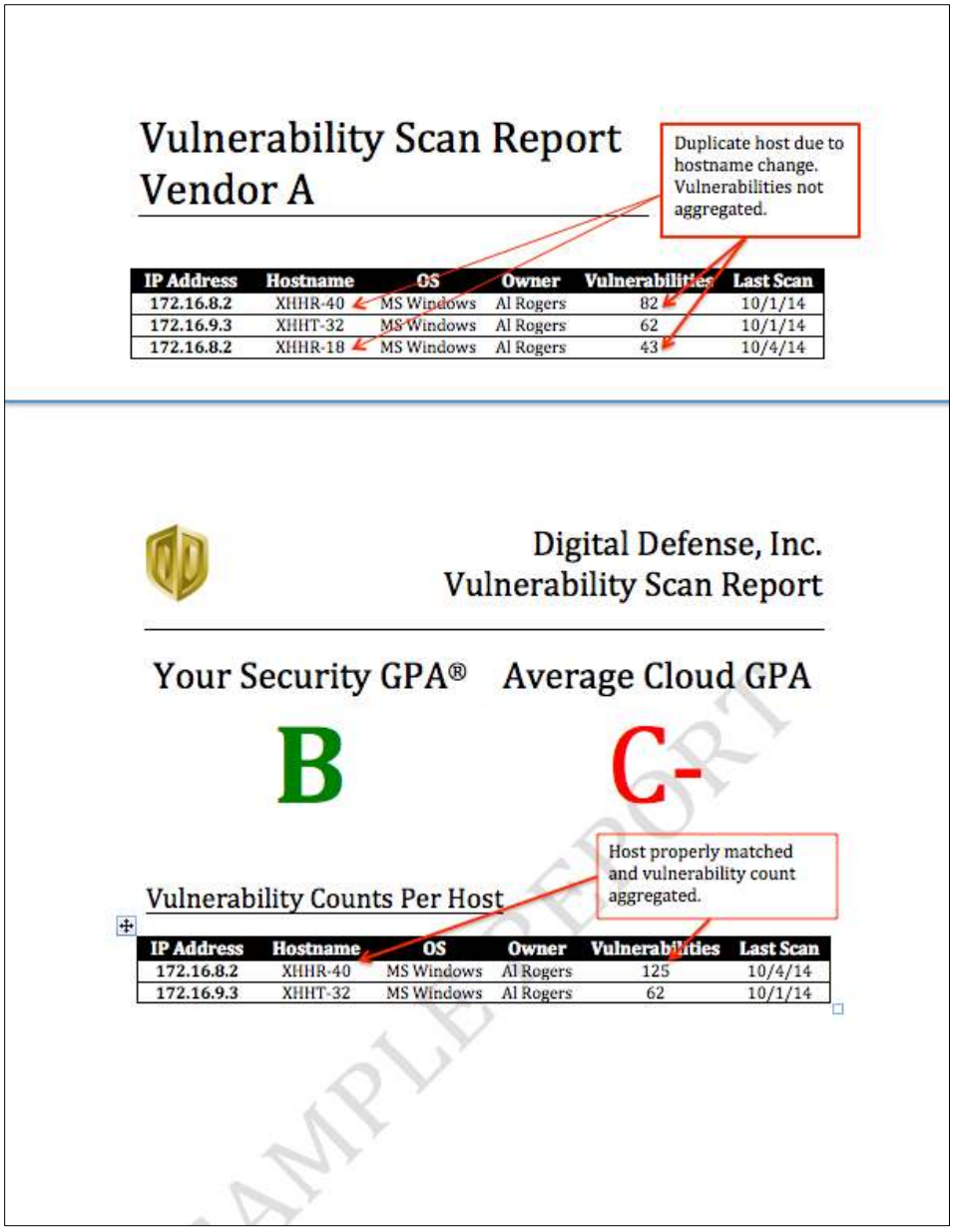
To provide a little more insight into this, let's look a little deeper into the specifics of a host's discoverable characteristics. These include IP address, various Hostnames such as DNS Hostname and NETBIOS Hostname, MAC addresses in some but not all cases, Operating System (with some level of accuracy), IP address of the parent router, and many other dynamic characteristics. Each of these host characteristics may change due to the nature of the host's configuration. For example, a laptop that is configured as using DHCP may change its IP address each time it joins the network. Even a Web Server or a printer may be assigned a different IP address as compared to a past configuration, simply as a result of regular ongoing IT administration. Hostnames may change due to a variety of reasons, such as IT administration activities related to aligning names with newly adopted naming conventions. The reasons for changes to these characteristics are many and varied.

I have described this challenge numerous times to a wide range of audiences, and occasionally have come across some skeptics and critics. Skeptics express their disbelief and argue that VM products must somehow solve this because the VM industry is very mature. In response to this, I share that though the industry is now mature, it has evolved from one where during the birth of the industry, VM providers were more focused on single point-in-time assessments and not as focused on managing historical findings, and as such had not invested in adequate host reconciliation technology. Further, for many years, clients were often too overwhelmed with the findings of even one vulnerability assessment, and therefore did not investigate the consequences of the time related aspects of their VM solution. Clients have now started to concern themselves about this crucial VM need, and even though the industry has evolved to one in which VM providers are touting the advantages of continuous assessments, most have not evolved their host reconciliation technology and as a result, struggle to provide effective and accurate VM time related capabilities. In terms of critics, I have met several who claim the challenge is quite simple to solve. Do not be fooled by such claims, as this is clearly not the case. If the solution were this simple, we would not be seeing the related posts found on the public vendor community support boards of the most recognized names in VM solutions[1].

One recent critic claimed the MAC address of a host never changes. This is simply not true. First, a MAC address is always discoverable IF a scanner is located on the same network segment as the host being scanned; unless one deploys a scanner on every network segment, the MAC address is not necessarily detectable. Secondly, failure and replacement of a host's NIC card (a common occurrence) results in a different MAC address for the host. While this hardware failure scenario does not occur all that frequently, it was identified by an IT administrator during a recent interview conducted for the purpose of a related study involving the nature of host configuration changes, and is covered further in this whitepaper. It is for this very reason that none of the major vulnerability assessment vendors use MAC address exclusively for tracking hosts to real world assets. In fact, the largest vendor in the industry does not use MAC address at all for host tracking.

The reality of this challenge is such that there is no unique characteristic, nor any perfect, simple set of two or even three characteristics that may be used to solve this challenge, which I refer to as the "Host Reconciliation challenge."

Having covered the above, it is important to note that network assets deployed within virtual environments of cloud providers are often built and managed within the given cloud service provider's platform. Many of these offerings assign a unique identifier for each virtual asset built and allow one to interface with the cloud provider platform in order to query information related to the virtual asset. Therefore a VM solution which interfaces with such virtual environments may easily use the virtual asset's unique identifier, which is obtained via the cloud provider API, as the matching key in order to perfectly match the asset as it appears within its point-in-time assessments. Having noted this, unless an organization has completely virtualized all of its endpoints, their VM is subject to the host reconciliation challenges covered in this whitepaper.



Example of Incorrect and Correct Host & Vulnerability Reporting Over Time

Study on Prevalence of Changes in Host Characteristics

Digital Defense, Inc. (DDI) recently conducted a study on the prevalence of changes in host characteristics, aimed at gaining further insight into the accuracy of our own host reconciliation matching algorithm. We performed the study using asset and assessment data collected from our cloud-based VM ecosystem. The findings revealed that our host reconciliation algorithm is highly accurate. In addition, we were surprised by the significant volume of host characteristics changes the study revealed.

DDI Reconciliation Algorithm Background

The DDI Frontline™ Solutions Platform (Frontline) Vulnerability Management (VM) system offers two types of assessment scanning methods; remote network discovery and authenticated discovery. When authenticated discovery scanning is used to perform assessments, Frontline is able to uniquely identify the scanned hosts and therefore map these to real world assets with 100% accuracy. When using remote discovery scanning, Frontline employs a patent-pending algorithm to map the hosts within an assessment to assets it had discovered during past assessments which represent an organization's real world network assets. Analogous to methods widely used to match scanned fingerprints to previously stored fingerprints in a database, the Frontline Network Host Reconciliation algorithm compares over 20 remotely-discoverable host characteristics for each assessed host, and compares these with the characteristics previously discovered and stored for the organization's real world assets.

Validating Accuracy of Reconciliation Algorithm

In order to determine the Frontline Reconciliation algorithm accuracy level, we considered a large set of assessments previously performed within the DDI cloud. These assessments included only those for which both remote and authenticated methods were used within the same assessment pass. Because these included the authenticated discovery method, matching discovered hosts to their real world asset counterpart is 100% accurate. The hosts within these assessments were then passed through the reconciliation algorithm to determine what they would have matched to if only the remote discovery method had been used.

Based on this part of the study, the Frontline Reconciliation algorithm was found to be 99.7% accurate in matching assessed hosts to their real world assets.

Determining Degree of Host Characteristic Change

In this part of the study, the goal was to determine how often host characteristics (especially those which are typically used by VM vendor products to match assessed hosts versus real world assets) change over time.

The logistics of this study were as follows:

- The analysis included only those assets which satisfied the following criteria:
    - Assets which were matched to hosts whose IP address fell within recurring assessment scan ranges, and for which were scanned at least 10 times across a 6 month period.
    - Included assets scanned using remote network discovery.

- Host Characteristics examined included:
    - IP Address
    - DNS Hostname
    - NETBIOS Hostname
    - Host Type - Client machine, Server machine, Printer, Firewall, Domain Controller, Device
    - Operating System
    - MAC Address
- Data was aggregated in several different ways as follows:
    - By Host Type
    - By Operating System
    - By Scan Type - Internal scans separately from External Scans

The findings from this study were quite voluminous, so I will focus on only a cross section of the results in this paper, specifically findings related to server type machines and client type machines. The following table summarizes the change rate findings for devices deemed Server machines:

Host Characteristic	**% Change**
IP Address	4%
DNS Hostname	46%
NETBIOS Hostname	34%

Table 1 Server Host Characteristic Change Over Time

The study revealed that, over a 3-month period, an average of 4% of the real world Server machines had a change in IP Address (relatively low, but significant in large networks),but 46% of these machines had a change in DNS Hostname change and 34% had a change in NETBIOS Hostname change (both extremely high).

A summary of the findings for devices which were deemed client machines appears in the table below:

Host Characteristic	**% Change**
IP Address	36%
DNS Hostname	42%
NETBIOS Hostname	20%

Table 2 Client Host Characteristic Change Over Time

Again the figures shown indicate the percentage of assets deemed client machines which experienced a change in the given host characteristic over the course of a 3-month timeframe. We see client machines have a much higher incidence of the IP Address characteristic than that seen for servers. (This intuitively makes sense given that client machines, which are used by humans to perform their job function, are often mobile in nature, e.g. laptops, and as such organizations often configure these to obtain their IP Address dynamically by way of DHCP.) The percentage of change in DNS and NETBIOS Hostnames runs high, as did the statistics for servers.

If a VM solution relies on these characteristics to match scanned results to real world assets and/or match scan results to previous assessments, the above results are quite alarming and clearly have serious implications. Assume for the moment that the large sample of networks and devices included in our study is representative of all networks (which is a fair assumption given the large sample size we used). One observation is that a VM system which makes the assumption that IP Address does not change over time for servers (and therefore uses IP Address to match assessed hosts to their real world asset counterpart), will not correctly match the assessed host to its real world asset 4% of the time across a period of 3 months. One may argue that 4% is a low number, but to put this into perspective, **an Enterprise organization having 100,000 Servers and using a VM system which matches in this fashion would have a resulting 4,000 of these servers with out-of-date and/or entirely inaccurate findings after 3 months.** As time progresses, more hosts that were not originally mismatched within the first 3 months of assessments could later also experience mismatches. Additionally, the problem compounds itself as time progresses because accurate matching requires an accurate basis; because of past mismatches, the basis is incorrect. This issue is even worse for cases where matching makes sole use of the other characteristics mentioned, since the prevalence of change is even higher than for the IP Address.

An Enterprise organization having 100,000 Servers could have 4,000 out-of-date servers and/or entirely inaccurate findings after three (3) months.

Given the much higher churn rate seen among client machines, scanning results comparisons can become virtually useless in a very short timeframe.

Why Hosts Change Over Time

Given the very high incidence of host characteristic change our study revealed, and given the ramifications of these statistics on the validity of results comparisons, we wanted to understand the nature of these changes to host. We held many interviews with IT Administrators at various companies to get a sense as to why these change are occurring within their organization. Some of the prevalent reasons are as follows:

- Change to IP address for servers which experienced a failure
- Change to IP address for servers related to IP segment changes
- Change to Hostnames due to adoption of a new naming convention
- Change to Hostnames after discovering a given name was not in accordance with their current standard naming convention

There were many other reasons provided by IT administrators explaining the host characteristic changes revealed in the study. Further, none of the individuals we interviewed were surprised by these findings; on the contrary, most indicated they make these configuration changes on a daily basis in the normal course of business.

Various Host Tracking Algorithms Used by Major Vendors

We performed some reconnaissance to determine how some of the other major VM vendors solve the challenge inherent with using the VM remote network discovery scanning method. With the exclusion of DDI's Frontline VM system (for which I provided insight into the algorithm used), we found the majority of vulnerability assessment vendors in today's marketplace employ very limited host tracking capabilities. I share two different methods which are used by two of the largest VM solution vendors, both of whom continue to receive the highest ranking possible by widely-respected market analysts.

Single Selectable Tracking Key of 3 Possible Characteristics

In this method, the product matches an assessed host to its real world asset by using a single host characteristic referred to as a "Host Tracking Key." The Host Tracking Key is one of the following: IP Address, DNS Hostname or NETBIOS Hostname. By default, the product uses IP Address as the matching key for all assessed ranges. However, the Host Matching Key is configurable by an administrator-level user who can specify different key method choices for hosts scanned within different specified IP Address ranges. At any time, the user may change the key method for a given IP address range, and this change results in a re-mapping of discovered hosts to real world assets.

For example, assume a company uses a VM solution which employs this method, and they quickly ramp up and start assessing on an ongoing basis. By default, IP Address is used to match assessed hosts to the organization's real world assets. If the organization eventually notices that many of their hosts are being mismatched to their assets over time, especially for cases where hosts obtain their IP Address by way of DHCP, a new matching key method such as NETBIOS Hostname could then be configured as the matching key within the solution only for the ranges where DHCP is in force. Once the new key method has been specified, the product re-maps all previous applicable assessed hosts to real world assets.

Although this method is flexible and allows a user to control host tracking to fit various network characteristics, it assumes the user understands the matching challenge and will detect mismatches. (Since many users are not aware of this problem, this assumption is seriously flawed.) Further, even though it is flexible, its matching capability is immature and, as we have seen in the study, is inadequate to handle the normal degree of host characteristic changes that occur within large user organizations' networks.

Multi-Key/Multi-Key Conflict Resolution – 3 Characteristics

This algorithm matches first with priority given to assessed hosts that match to assets on the basis of both of two keys - IP Address and Hostname. The un-matched assessed hosts remaining after this first pass are then first matched using the Hostname if it is a present characteristic for the asset. Otherwise, the IP Address is used. This method differs from the previous method in that it is entirely hands-off and requires no intervention. Unfortunately, this also means one can't make corrections to resulting mismatches.

As compared to matching on a single characteristic, this method does give higher confidence levels for matches involving those assessed hosts and assets which have the two matching characteristics in common. Unfortunately, since our study shows that both keys experience high rates of churn, the potential for erroneous "matches" remains high. Hence, the results of our study show this method to

be woefully inadequate. Additionally, if erroneous matches are discovered, there is no way to apply corrections to the mappings. Errors remain errors, period. As a result, the historical asset finding information quickly becomes unreliable and unusable.

Consequences of Host Tracking Errors

In this part of the whitepaper, I describe in greater detail the different types of possible mismatch errors resulting from the inadequate matching methods employed in most solutions on the market today with respect to matching the point-in-time assessed hosts to their real world assets. I also describe the risk related consequences of these matching errors to an organization.

Types of Mismatches

There are three different types of mismatch errors that result from the inadequate historical tracking methods used by most VM solutions with the primary and most widely-used scanning method today, namely remote network discovery scanning. These are:

- *Unmatched and Excluded*
  When an assessed host is not matched to any asset, it is left unmatched to any real world asset. Further, the unmatched asset is entirely excluded from the results of the assessment. (While I appreciate that this sounds ludicrous, I assure you it is true. It also happens to be the method employed by the largest VM solution provider at present.) Although the result is clearly undesirable and there are clearly unfortunate consequences, there is a logical and valid rationale behind the method. The solution first employs an asset "mapping" process. The mapped hosts are then deemed to be the real world assets. As assessments are performed, any host within an assessment which is not successfully matched to a real world asset is then deemed an asset which has not yet been mapped, and it is assumed the asset will be picked up in a subsequent mapping scan. Exclusion occurs when the host tracking method selected is not the IP Address method, but instead one of the Hostnames (DNS or NETBIOS), and where the Hostname for the assessed host does not match to a Hostname for any of the previously-mapped assets.

- *Unmatched and Added as Duplicate*
  Most vendors do not employ the above described "mapping" process regimen. When an assessed host should have been matched to a real world asset but instead remains unmatched, it is deemed to be a new real world asset and is, as a result, added as a new asset to the list of real world assets. Effectively this adds a duplicate asset to the list of real world assets. This situation is illustrated in the following figure.

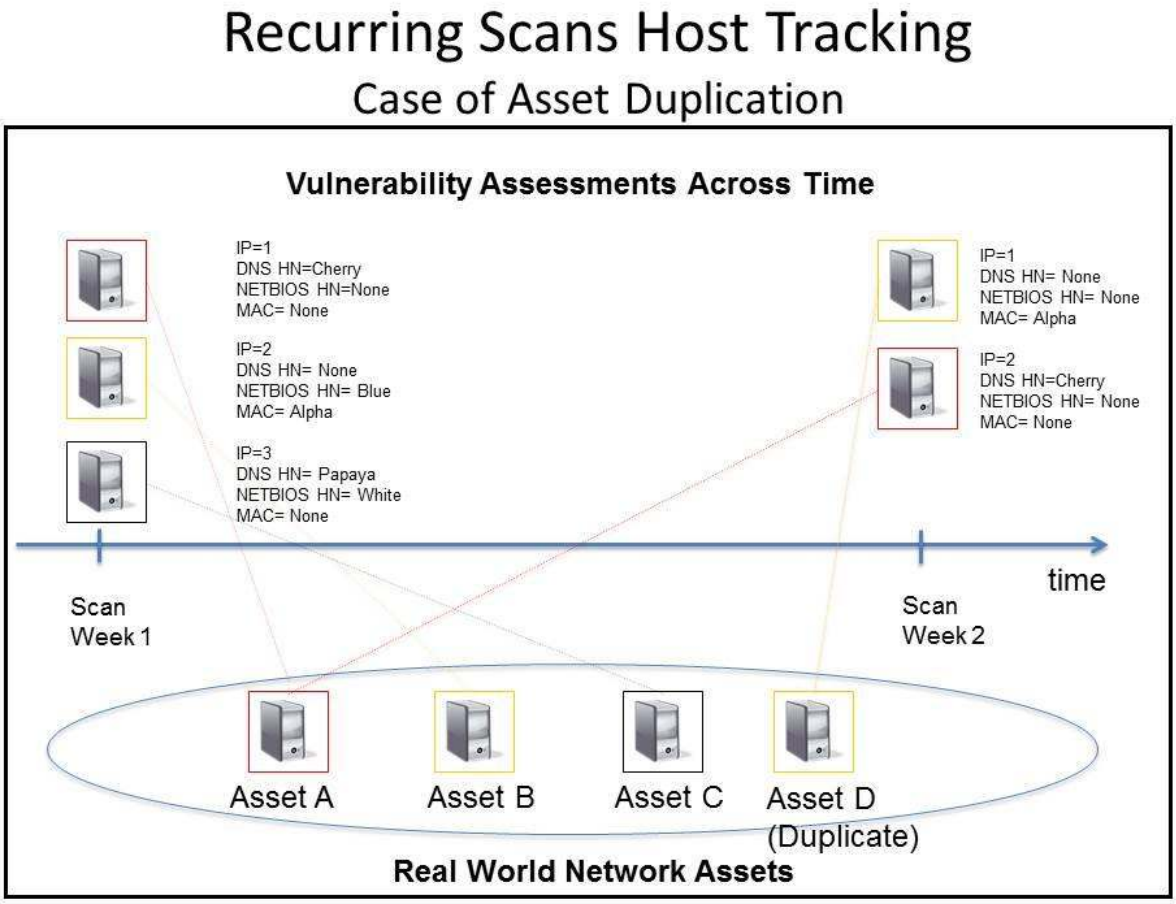


Figure 3 Mismatch case of Asset Duplication

- *Mismatch to incorrect asset*
  In this case, an assessed host is incorrectly matched to an existing real world asset. This can occur for a variety of reasons. In the context of the previously-mentioned methods used by major vendors, if an asset's IP Address for a Server was changed due to normal IT administration and the new IP Address assigned was at some point used by a different asset that had been previously assessed, this unfortunate mismatch error occurs. The VM solution consequently produces "apples to oranges" results when comparing current scan data to historical results. This situation is shown in the following figure.

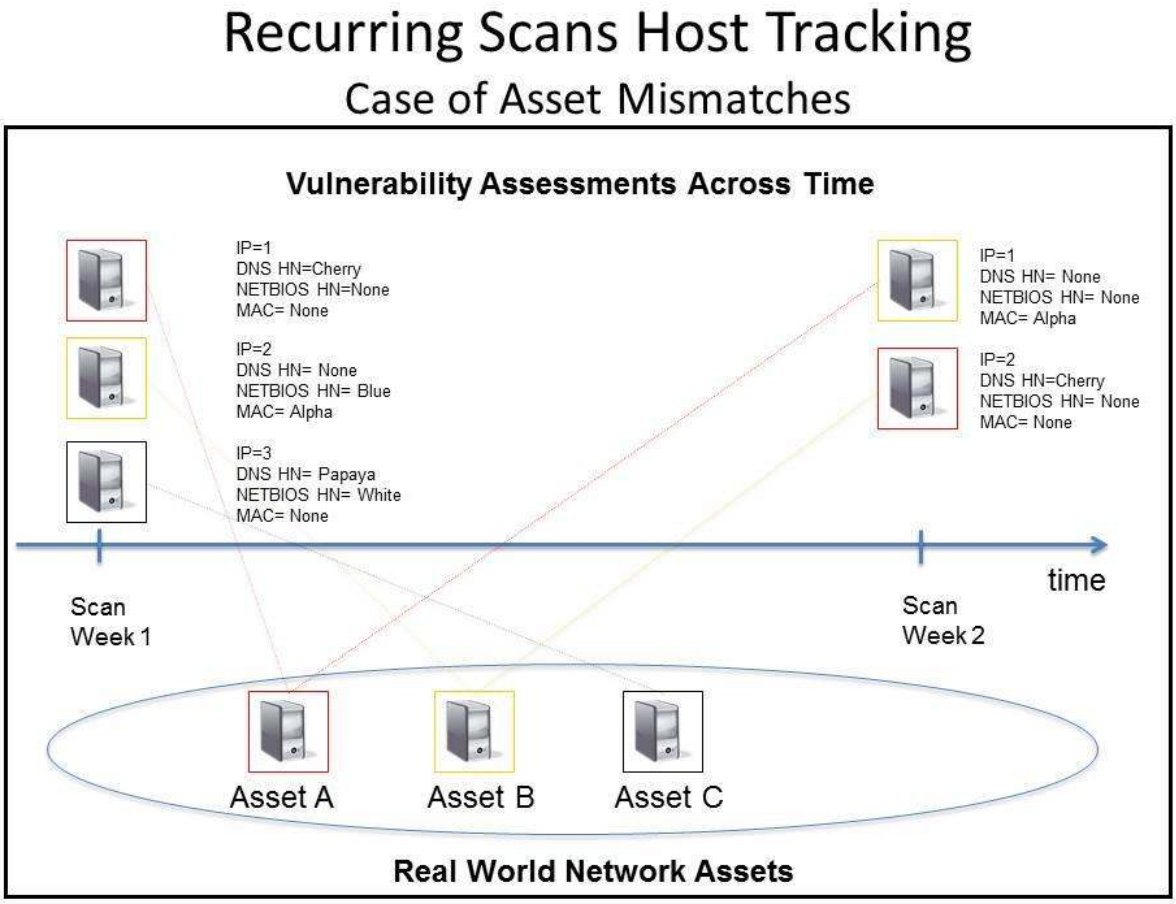


Figure 4 Mismatch case of Assessed Host Mismatch to incorrect Asset

Consequences of Mismatches

With the above types of mismatch errors in mind, the consequences of such errors are:

- When a host is dropped from the results of an assessment because it does not match to a previously "mapped" asset, that asset's findings are not updated. If this occurs on all subsequent assessments for the given asset (quite possible), in effect the asset is never re-assessed and the information related to that asset grows stale and becomes irrelevant. This has a direct consequence to the risk gauge for that asset, and impacts an organization's ability to make appropriate decisions and properly assess exposure to cyber-attack risks. Further, any integration of the VM product using information related to this asset is in effect drawing conclusions which are uncertain at best. The worst part of this error is that the client is in most cases unaware of the issue; when the host is dropped from the assessment as a result of the mismatch, there is no alerting mechanism available to identify/report the mismatch.

- When a host is added as a new asset as in the second mismatch error described, the user organization's information is impacted in several ways. First, the organization must spend time investigating where the asset is located, who is responsible for the asset and so on. In some

cases, vulnerabilities for the non-existing asset are assigned to individuals who then proceed to investigate possible remediation efforts only to later discover the asset was a duplicate asset (creating duplicate effort). Finally, the organization is provided with misleading information within their cyber-risk reporting due to the presence of many duplicate assets cluttering their asset view.

- One of the worst and most problematic mismatch errors is the third one previously described, where an assessed host is mismatched to an incorrect real world asset. When this happens, typically the vulnerabilities found within the assessed hosts are different than those actually present on the incorrectly matched real world asset. As a result, vulnerabilities on the real world asset are declared "fixed" and new ones are found to exist on the "new" asset. This is another instance where IT resources are wasted, since the vulnerabilities on the duplicate asset are assigned to staff for investigation/remediation when in fact none may actually exist on the given asset. Another consequence of this mismatch error occurs when enforcement technology integrated with the VM systems concludes vulnerabilities are fixed, when in fact they may actually still be present on that asset. Depending upon the value of the integration and the protection provided by the enforcement technology, the organization may be left in a state of high risk while being told "all is well" by the VM solution.

Conclusion

The accuracy of one single assessment is very important and remains a benchmark traditionally used to compare various VM solutions and products. The truth is that, in today's mature VM solution market, scan accuracy does not vary significantly from one vendor to another. Unfortunately, far less attention has been paid to the accuracy of these results as they relate to comparison of scan results over time. The reality is that the findings portrayed within the "asset views" of the VM systems used by most organizations (including the Fortune 500 Enterprises) are far less accurate than we once believed. Organizations use this flawed information to guide their security decisions, and integrate it with their security enforcement technologies. The reason this data is flawed is the inadequate techniques used by most VM solutions in dealing with the crucial matching challenge as described extensively in this paper. The results of the Network Host Reconciliation study summarized in this paper show clearly that hosts which are part of organizations' networks are actually subject to significant and frequent, illustrating the potentially severe negative impact these inadequate techniques introduce, seriously impairing the usefulness of the data presented to the users of these solutions.

The primitive algorithms found within the inner working of the VM solutions supplied by even the largest of the vendors in the space are seriously flawed, and cannot correctly track the findings for the ongoing assessed hosts in the presence of the dynamic change our study exposes. As a result, an organization using such solutions must take extreme care not be fooled by the risk profile portrayed by these products, and instead must question the matching technology used within them and take action to avoid the pitfalls the present (mainly a false sense of security or the chasing of phantom problems). Organizations who find their current solution is inadequate and cannot/will not solve this challenge should either procure a replacement solution that employs more robust reconciliation methodology or, alternatively, integrate with a third party solution which can detect and eliminate the errors present in the information provided by their current VM solution. Users who do neither will, unfortunately, continue to be provided with "information" riddled with erroneous data, with no way to separate the truth from the fiction.

References

[1] Public internet community posts related to vulnerability management challenges available upon request