A TagVault.org White Paper

TagVault.org
c/o IEEE-ISTO
445 Hoes Lane
Piscataway, NJ 08854
732.562.6031
www.tagvault.org

The Future of Software Asset Management

By Steve Klos
Executive Director, TagVault.org

August 2009

This page intentionally left blank

Contents

Introduction

Every organization that purchases and uses software has multiple reasons for knowing exactly what software is installed in their infrastructure and how best to manage it. The reasons fall into three major categories:

- **License Management** – accurate software identification is a critical input to software asset management (SAM) and license entitlement reconciliation.

- **Security Rationalization** – knowing exactly which software titles are installed, who created the software, and which files are related to the software titles provides a complete security posture for the organization.

- **Organizational Infrastructure** – providing techniques that organizations can use to track software testing and rollout procedures, as well as purchasing and usage patterns, gives companies information they need to more effectively distribute and utilize their software assets.

These issues have a direct and immediate impact on software purchasers, but they also have an impact on software publishers and SAM tool and service providers.

This white paper explores some of the reasons that SAM programs are currently so difficult to implement, and what the market is doing to make it cheaper, easier and faster. It does not go into details about implementing an effective SAM program, as there is already a significant amount of valuable content available. (One source is the test preparation guide written by Agnitio Advisors for Microsoft's SAM certification test (70-673). Another source is the SAM Guides website.)

License Reconciliation

Organizations implement SAM processes to ensure licenses are reconciled for a number of reasons. Often, these processes are put in place to ensure that an internal, publisher or trade group audit of the organization does not find over-deployment of software. Reconciliations may also be instituted in order to eliminate over-purchasing of software, provide charge back reports or simply to validate that organizational policies are adhered to.

When doing license reconciliation, software identification of all titles on all corporate computing devices is a key input. Unfortunately, the reliability and consistency of current software identification procedures leaves a lot to be desired.

Software Identification – State of the Art?

There are a large number of tools on the market that provide some form of software recognition functionality. These tools vary in breadth and scope of the software they recognize as well as in the platforms they support. In general, these tools utilize a signature library that is used to reverse engineer inventory information to try to determine which software titles are installed on a particular device. The problem is that there are too many software configurations these tools cannot identify properly, either because the inventory does not provide enough information, because the software products themselves have multiple configurations that cannot be identified by an inventory process, or simply due to lack of tool support for a particular platform. Additionally, different recognition tools identify software differently, meaning the information from one tool is often difficult or impossible to use with another tool.

Howard Hastings, ITAM Evangelist for CA, has written a white paper identifying many of the issues with current software identification practice. His white paper, *Why Software License Management Is so Difficult – and How to Simplify It*, is available as a link from the TagVault.org web site (found in the content library – you must create a free account to access this area of the website).

This paper won't cover why software identification is difficult –Howard's white paper has already covered many of these issues. Instead, we will review what the market is doing to resolve this problem for all SAM eco-system members.
Let's use a simple and very common example. A software discovery process may find the following:

Filename: photoshop.exe
File size: 17,956,864
File date: Oct 17, 2003, 2:50 AM
MD5: dbfbb5e93390d2f57da4669bb117e7a1

A software recognition process may identify this application file as the following:

- Photoshop Professional – **or** – Photoshop Extended – **or** – Photoshop CS4
- *conditional* – May be part of a suite, or not
- *conditional* – May be an upgrade version, or not
- *conditional* – May be a trial version, or not

Add to this one more variable – which is that one discovery system may identify the application with one name, while another identifies the application with a slightly different name (for example Photoshop Extended vs. Adobe Photoshop Extended). This inconsistency means that consolidating data between the tools requires additional normalization efforts.

That leaves a lot of open questions for the SAM practitioner to deal with. Now multiply that problem times tens of thousands of computing devices, across hundreds of software publishers, thousands of software titles, numerous bundling and OEM options, and multiple platforms – the challenges add up very quickly!

Resolving Software Identification Problems

Software installed on any platform should be easy to identify. After all, when software publishers decide to audit an organization, they know exactly what to look for. In the case of the Photoshop example above, Adobe generally accesses the Adobe product key to identify the product – unfortunately, this information is only available to Adobe. If only that specialized knowledge, unique to each publisher, could be provided in some structured fashion to the software purchaser. In fact, this information *can* be provided by the software publisher in a very inexpensive and effective manner, by using software identification tags as defined in the

ISO/IEC 19770-2 standard. This standard provides a structured set of data that can be used to uniquely identify software regardless of the platform, from an iPhone to the largest mainframe.

The ISO/IEC 19770-2 standard defines an XML file that contains seven mandatory data elements that uniquely identify the publisher and application details. The standard also defines 30 optional elements that provide significantly more contextual information to the SAM practitioner. Finally, the standard allows the XML file to be extended with additional information as required.

"Instantly, the SAM practitioner knows exactly which version of an application is installed and, if installed as part of a suite, which other software was also installed as part of that suite."

Instead of having to deal with the software identification issues outlined in the previous section, software ID tags will provide comprehensive details directly from a discovery process (regardless of the tool used). For example:

- 1,000 copies of Adobe CS4 Web Premium Version 4.0.0.0 Volume Version installed
    - 1,000 copies of Photoshop CS4 Extended, Version 11.0.0.0
    - 830 copies of Acrobat 9 Professional, Version 9.0.0.0
    - 200 copies of Dreamweaver CS4 installed, Version 11.0.0.0

Instantly, the SAM practitioner knows exactly which version of an application is installed and, if installed as part of a suite, which other software was also installed as part of that suite. Not only that, but the software identification tag may also specify that there are 20 trial copies of the software installed in the organization.

This information provides the exacting detail required for a SAM tool to automate the license reconciliation process. Additionally, the data is the same regardless of the discovery tool used, meaning consolidation of data from multiple tools can be completed easily. The ISO/IEC 19770-2 standard ensures software purchasers get exactly the information they need to manage their reconciliation process.

Significantly more benefits are seen when TagVault.org-certified tags are provided by the software publisher. TagVault.org ensures that the tag contents conform to the 19770-2 specification, and digitally signs elements in the tag to ensure the data is not modified by any user. TagVault.org also ensures the use of standardized terms for information such as the publisher identifier, media used for the installation, target customer, whether or not the software is a trial version and so on. Imagine a SAM practitioner being able to identify all trial versions of all software products currently installed in their environment by simply filtering a report on one pre-defined term. This type of reporting functionality is unheard of when current software recognition libraries are used, but will become common as certified software ID tags are included in software products.

How Can Software ID Tags Make a Difference Today?

Certified software ID tags that are included with new software titles will take some time to roll out. Organizations first need to integrate tools and processes to create the software ID tags – and TagVault.org will obviously help in that area.

However, TagVault.org-certified tags are not only useful for new software products. TagVault.org provides a repository of software ID tags, enabling community- and publisher-defined tags for legacy applications to be used as signature files to identify software. Software ID tags can be defined for any kind of software running on any platform – commercial applications, in-house developed applications and applications installed on an operating system such as a UNIX server can all have tags defined using the same method. Once defined and stored in the TagVault.org repository, the software tag can be used to identify installed software anywhere. Although the identification still relies on a signature process, the signature is at least common to all discovery tools and is created once, and is then available to current and future discovery tools that access the tag repository.

Now, when a software ID tag is created by anyone and checked into the community area of the TagVault.org repository, that tag can be used by other companies that also use that software title. Creating a community-accessible repository allows the collective knowledge of the full membership to significantly reduce the workload of individual SAM practitioners. This community-supported repository will also be available to software tool providers that are members of TagVault.org and that provide discovery and SAM tools to the market.

Collecting Software ID Tags

Now that the market has software products (such as the entire Adobe CS4 product family) that include software ID tags, and there is an organization to support a community-based tag registry as well, what next? How can the information be used?

A number of SAM tools can access information from structured XML files today. Altiris has published an article detailing how to pull information from a software ID tag and store it in the Altiris CMDB. SMS has a similar capability, as do other SAM and desktop management tools. Since the complete Adobe CS4 product family now includes software ID tags, all discovery and SAM tools need to recognize software ID tags as well, or be left behind by the competition. This ensures that the SAM practitioner will get accurate and, as importantly, consistent information from any tool they use for discovery or SAM processes.

What About Entitlements?

Software ID tags are defined by the ISO/IEC 19770-2 standard, which should be published in 2009. Following the software ID tag standard is the software entitlement tag standard – ISO/IEC 19770-3. This standard is closely associated with 19770-2, with a number of shared elements that are specifically designed to automate software entitlement reconciliation.

Today, software entitlements are specified in the software license agreements an organization purchases. These are legal documents that specify the terms of use for a software license. Unfortunately, these documents are often difficult to interpret and they provide no specific guidance on how to measure and track software utilization to compare against the entitlement. In fact, there are many entitlements that are specified but cannot be tracked – during an audit, these often end up with an agreement or an estimate of the quantities.

The 19770-3 draft standard is designed to provide a structured definition of the specific metrics that must be measured on various computing devices in the organization to determine entitlement utilization. Depending on the license type, this may take the form of a tag that specifies the following (note, these are not specific elements from the19770-3 draft, they are English language examples):

- Product ID: sample-product (from SWID tag)
- Track On: local device
- Tracking: installation
- Entitlements Available: 6000

This type of structure would be used to track software entitlements based on the software being installed. There are other structures that support subscription licenses, client access licenses, usage licenses, per-processor licenses, and more.

Key to the 19770-3 draft is the ability to specify where an organization (or tool) needs to track the information (on the device itself, on a server, on a virtual server, etc) and what specific metrics need to be tracked (an installation, network access to a specified port, process execution table, etc).

With ISO/IEC 19770-3 software entitlement tags provided to purchasing departments, and software that's already using ISO/IEC 19770-2 software ID tags, the process of reconciliation will be significantly more automated than it is today. The ISO/IEC 19770-3 standard also prescribes a methodology for software publishers to specify exactly what a software customer needs to track. This level of detail is already being requested by larger software customers, but the software entitlement standard will make it available to every customer in a manner that SAM tools can use for automated reconciliation.

19770-2 and 19770-3 Process Flows

From a process perspective, it's important to note that a software ID tag follows the software – when a software

product is installed (or in the case of a SaaS title, when the software is used). This means that the software ID tag is consistent for a specific installation routine distributed by a software publisher (obviously, there will be unique items in a any specific installed software ID tag – such as the serial number, activation status and other elements unique to that particular installation). The software entitlement tag, on the other hand, follows the purchasing process, and will be generated uniquely for the specific customer at the time of purchase. This means that negotiated terms will be included in the tag without additional effort from either the software publisher or the software customer. The process will be as automated as the generation of the software entitlement paperwork is today.

Security Rationalization

In organizations today, software inventory details are often not considered reliable enough to assist with security assessments. This is due to the lack of accuracy and comprehensiveness of the software inventory tools.

> "With certified software ID tags, software inventories can authoritatively identify software titles, and all associated files that are part of that application can be filtered out of the inventory as being 'known' files."

Security in Depth

Discovery tools often provide some level of software identification, but do not validate all the ancillary files for a specific title. These ancillary or helper files are often ignored due to the scale of the issue (a reasonably loaded notebook computer can have upwards of 150,000 files). This means that the underlying captured inventory will have numerous files that could be executed but which are left unrecognized. These unrecognized files may not cause problems in an audit, but they could be malicious applications that either mimic a known application filename, or that simply look like "helper" files (for example, a file called common.exe could be a helper file, or it could be a virus or malware program).

Organizations can and do rely on virus scanners to minimize the impacts of this type of file, but what if that file was created specifically to capture internal data, and was not identified as a threat by virus scanners? In depth software identification, including recognizing which files are associated with an installed application, allows organizations to apply known and valid file filters to their discovered inventory to ensure that only the software that's expected to be installed on the organizations computers is, in fact, installed.

Inventory Support for Software Security

A complete, accurate and trusted software inventory can provide a significant amount of data to improve the security posture of an organization. Because software inventory is required for many other organizational processes, including SAM processes, if done properly, there will be little or no additional cost involved in creating an inventory that can be useful for security audits.

Software ID tags by themselves do not provide the type of provenance information necessary to authoritatively identify a software title, or the associated files for that title. To provide this type of support, the software ID tag must be digitally signed by a known and authenticated source, with the signature including the package footprint (which is the list of all associated files and other components published for this title). Software ID tags registered with TagVault.org can have the package footprint element digitally signed by TagVault.org, and the software ID tag can then reliably be used to identify not only that the software was created by a specified publisher, but that the files associated with the software title were also created by that publisher.

Note that this is not based just on the filename, but also on a secure digital fingerprint of the file that can be used to authoritatively identify whether the file has been modified in any way. The resulting unidentified files that remain in the organization's inventory will either be user data files, or unknown application files that may require further analysis.

Managing Exceptions

Using the community-supported repository provided by TagVault.org, organizations can choose which software tag creators they choose to trust, if any. They may choose to trust only software tags created by their own organization or only tags created by members at a certain level of membership within TagVault.org.

Community-supported software tags allow the organization to create its own software ID tags to filter out "known" files. Over a very short period of time, organizations can create software ID tags for software that is authorized for use on the computing devices in the organization's network. As future inventories are processed and filtered, any outstanding software files – especially executable files – that are not filtered out are considered exceptions that the desktop management or security groups can readily identify and analyze as appropriate.

Organizational Requirements

Organizations that buy and use software need to specify policies for what software may be used and how it can be used. This is important not just to ensure the organization is complying with software entitlements, but also for the security of company data and management of potential corporate liability.

Doc Burnham has addressed the benefits of end-users using software tags in a white paper entitled, *"Using ISO/IEC 19770-2 Software Identification Tags to Enhance Software Asset Management"* – available from the TagVault.org web site (found in the content library – you must create a free account to access this area of the website). This white paper expands on some of the topics Doc raises in his white paper.

There are many freeware and community-developed software titles that are free to download and install. Many of these titles can be very useful to the individual, or the organization. However, there are numerous titles that open security holes in the organization's defenses. For example, if LimeWire, BitTorrent or any of the Internet file sharing utilities are installed on a system and configured to share files that have confidential information, the organization may be providing confidential information directly to its competitors. Organizations need to specify if these types of utilities can be used, who is authorized to use them and for what purpose.

An organization also needs to specify that inappropriate content should not be installed on any of its computing devices. This policy may apply to explicit programs, as well as to music and video files that are not owned by the organization. Having these types of files installed on an organization's computing device can open the organization to liability issues.

Managing by Best Effort

Having set policies, how does an organization manage to that policy? Users can install software titles and change filenames or take other actions to mask the installation from inventory routines. Desktop management teams often apply a best effort approach to managing these types of issues. If they have not applied software tags to their internal software (or used software titles that are provided directly by software publishers), there is no way for the organization to manage by exception – there will be always be too many unknown files to deal with.

What Gets Measured Gets Managed

Turning the problem of organizational management around, if all titles in the organization are identified and the files associated with the titles are also identified, the organization can start to manage by exception – when an unknown software title or file is discovered, the desktop management, SAM or security teams can validate that the software is appropriate to be installed on the device.

Today, organizations cannot get to this point for a few reasons. First, having one group within the organization identify every software title and all supporting files is an onerous task. Since there is no industry-standard method available for general sharing of trusted identification information, each company is left to its own resources to identify every application and file. Even if the organization uses a commercial application recognition library, the library will never have a complete list of titles for many of the same reasons a company can never get a complete list – it's too large a job for one organization to manage on its own – it requires a community effort.

Secondly, if a company does go through the effort to create software identification details for one discovery engine, that data is not transferable to a different discovery engine. Many larger organizations have multiple discovery tools in different parts of the organization, so this transfer limitation reduces the usefulness of custom discovery information significantly. Additionally, the organization may also choose to replace one discovery engine with another,

which necessitates discarding all the custom efforts the organization has already made, and starting over with the new tool.

TagVault.org resolves both of these issues. First, by providing support for a community-based software ID tag repository, it provides data that can be used by any organization or tool developer that is a member of TagVault.org. Organizations can determine which community tags they want to use. They can limit use of the repository to only access certified ID tags and ID tags that they have created, or they may also choose to use ID tags that have been created by other corporate members. The ability to directly use ID tags created by other entities spreads the load of ID tag creation across a much larger population, and increases the number of titles that will be identified as "known", making management by exception a possibility.

Second, TagVault.org provides a common repository for standardized software ID tags. Once an organization adds its tags to the TagVault.org repository, those tags will be available to any and all tools that support software ID tags. Organizations can easily move their internally created ID tags to any other discovery or SAM tool that they choose to use.

TagVault.org provides the first economical process that makes management by exception for all software titles in the organization a possibility!

ISO/IEC 19770-2 Software ID Tag Standard

The ISO/IEC 19770-2 software ID tag standard is expected to be published in 2009. As of August 2009, the document had been voted to become a Final Draft International Standard (FDIS), which means that the document will be voted on one more time by the various national standards bodies associated with the development of this standard. Based on the last vote the Final Committee Draft version received, the draft is expected to be voted into becoming a published standard. Technical content in the draft is not expected to change between the version submitted as an FDIS draft and the final published document.

The ISO/IEC 19770-2 standard is an excellent first step in standardizing identification information that is provided with a software installation package. When created, the standard was written in a way that a certification authority would not be explicitly required – this was done to allow for easier market adoption of the standard. However, there are numerous ways that a certification authority will benefit the market. These include the ability to validate that early market software tags comply with the standard. Additionally, there are numerous identity details (referred to in the standard as regids) that should to remain consistent, as well as terms that it would be beneficial to normalize. Finally, providing a community-based software ID tag repository allows for the distributed creation of software ID tags and the ease of portability of tags that software customers need.

TagVault.org is a non-profit organization committed to providing support for tag certification and the software ID tag repository as well as the tools, services and knowledge necessary for organizations to create and use tags in the most efficient manner possible. As a non-profit, TagVault.org will do this at the lowest cost, and also offers the SAM community the security of knowing that the data will not be incorporated into a commercial entity and potentially be removed from general market access.

"The next standard under development is the ISO/IEC 19770-3 software entitlement tag standard. The software entitlement tag standard uses a number of the same elements as the software ID tag standard. This ensures that the reconciliation process is easy to automate. "

ISO/IEC 19770-3 Software Entitlement Tag Standard

The next standard under development is the ISO/IEC 19770-3 software entitlement tag standard. The software entitlement tag standard uses a number of the same elements as the software ID tag standard. This ensures that the reconciliation process is easy to automate.

This standard's development is commencing under the leadership of John Tomeny of Sassafras Software. A number of the individuals who volunteered to work on ISO/IEC 19770-2 are also part of the working group

creating the ISO/IEC 19770-3 draft. Individuals who are interested in volunteering to help develop the ISO/IEC 19770-3 standard should refer to the introduction to the working group web page hosted on the Sassafras Software web site.

It is expected that the ISO/IEC 19770-3 draft standard will have to go through additional market validation steps to ensure that the largest number of entitlement options can be supported. TagVault.org will do what it can to support this process as the development of the standard moves forward.

Summary

Software identification is difficult to do well, creates incomplete and inaccurate reports, and is expensive and resource intensive for organizations to work with. Some organizations are negotiating specific details in their entitlement agreements that specify exactly what needs to be discovered for software to be determined to be installed and are managing their assets to those specific expectations. However, most organizations, though they can be held liable for over-deployment, have simply accepted the lack of detail and accuracy involved in software discovery as part of the market expectations. The ISO/IEC 19770-2 standard and TagVault.org exist to remove the unknowns of software identification and provide a more functionally capable process that increases accuracy, and decreases costs involved in discovery procedures. As the ISO/IEC 19770-3 gets closer to completion, TagVault.org will support this standard as well, since it will provide the entitlement data that – when combined with the software ID data – will automate almost all of the SAM reconciliation procedures.

Joining TagVault.org provides the tools, services and knowledge organizations need to fully utilize software ID tags, and will allow any of the SAM eco-system organizations to move from a process that's difficult and expensive to a more automated, accurate and complete process.

Resources

SAM process information:

- Microsoft SAM test preparation guide - (partner.microsoft.com/40092833)
- SAM Guides web page - (www.samguids.com)
- End-user use of software ID tags whitepaper (www.tagvault.org)

Standards web pages:

- 19770-3 Other Working Group – (www.sassafras.com/iso)

About TagVault.org

TagVault.org is the registration authority for ISO/IEC 19770-2 software identification tags (SWID tags). Formed as a non-profit member program of the IEEE Industry Standards and Technology Organization, TagVault.org provides a shared library of software tools and technical knowledge that decrease the costs involved in creating and managing software tags. TagVault.org's registration process ensures tags fully conform to the tagging standard while also ensuring normalized tag contents. TagVault.org also provides communication forums and information sharing resources among software publishers, tool providers and SAM practitioners. For more information, including the benefits of membership, please go to www.tagvault.org.

About IEEE-ISTO

IEEE-ISTO is the premier trusted partner of the global technology community for the development, adoption, and certification of industry standards. Its mission is to facilitate the life-cycle of industry standards development through a dedicated staff committed to offering vendor neutrality, quality support and member satisfaction. Fostering the market acceptance, adoption and implementation of standardized technologies, ISTO Programs span the spectrum of today's information and communications technologies. To find out more about ISTO, visit www.ieee-isto.org.

TagVault.org
c/o IEEE-ISTO
445 Hoes Lane
Piscataway, NJ 08854

Phone: 732.562.6031
URL: www.tagvault.org