Fundamentals of Software Audit Data Collection – Hardware Inventory

By Christopher Barnett
Scott & Scott LLP

In order to effectively manage their software usage and to mitigate compliance exposure, companies need to know how to gather and analyze information regarding their product usage. While some software products may have unique data-collection requirements that ordinarily would not be applicable for other products, usage levels for many products can be measured using a common set of reports that companies can prepare themselves to gather. The purpose of this and other posts in this series of blog entries is to give companies insight regarding how to gather those datasets and why that information is relevant to their licensing obligations.

The backbone of deployment data for many products is a list of all devices in an IT environment capable of running or interacting with a licensed software product. This inventory should include all physical and virtual workstations and servers. In many cases, it also should include devices like thin clients that may not be capable of running installed copies of software, but that can be used to access software deployed in server environments. The list also should include information regarding the make and model of each device as well as the make, model and quantity of processors and processor cores for physical servers and virtualization hosts. (Note that information regarding virtualization environments will be discussed in a separate entry in this series.) Finally, in order to validate the completeness of the principal hardware inventory, it almost always is a good idea to obtain a secondary inventory from a source like Active Directory or an antivirus solution.

Tools

Most companies use automated, software-based inventory tools in order to gather information regarding the hardware and software present in their IT environments. Using tools is not absolutely necessary – especially in very small environments – where a physical machine count, purchasing records, and change logs may provide sufficient information to allow a company to know what it owns. However, tool-based inventories almost always represent a best practice, since manually maintained reports should be audited against other data sources. Furthermore, it is typically much less practical to maintain a software inventory manually, and the tools used to gather software-installation data for an environment typically can provide all the information needed for a hardware inventory.

The tools used will vary depending on the types of software being run in the environment and on the way that the IT resources are configured. For example, an “agentless” tool – which consists of a centralized installation that uses communication protocols to remotely gather information from other machines on a network – may be the best choice for companies that have relatively simple networks with limited numbers of domain-administrator-level credentials that can be used to access all connected devices. Agentless tools usually are less expensive and easier to deploy than other kinds of solutions. **Microsoft’s MAP Toolkit** is a free, agentless tool and is one of the most widely used solutions for gathering information regarding Windows-based environments. **Spiceworks** is another, popular, agentless tool.

However, in some cases, an agentless tool may be a bad fit. Some environments may be based on global deployments residing within multiple different domains that do not share a common set of administrative credentials. Alternatively, it may be necessary for an environment to be highly segmented for security reasons, with that segmentation effectively preventing an agentless tool from gathering information from resources in other segments. In these circumstances, an “agent-based” tool – consisting of a centralized management console and multiple, small “agent” installations on each computer deployed within an environment – may be a better fit. In that scenario, the agents typically do a more thorough job of gathering information regarding the machines where they are installed and often can be used to remotely manage those machines in addition to simply gathering data. Furthermore, agent-based tools can take advantage of encryption technology and a wider variety of communication protocols to avoid the kinds of network-configuration obstacles that may prevent agentless tools from providing meaningful information. Unfortunately, all of that functionality comes at a cost, as agent-based tools often are much more expensive and much more difficult to successfully install and configure. **Ivanti IT Asset Management** (formerly known as LANDESK) is a very popular agent-based tool. IBM’s **License Metric Tool** is another well-known (perhaps infamous) agent-based tool that may be required for companies using certain kinds of IBM software in certain kinds of environments.

Some smaller companies may not want to invest the capital or resources to deploy an IT asset discovery tool but still may not want to rely on manually-maintained inventories. In those circumstances, it may be feasible to gather information using scripts or system queries that provide the same kind of information as automated tools, but via an ad hoc process. In Windows-based environments, **PowerShell** queries can deliver a wealth of information regarding machines residing with the same Active Directory architecture and also regarding the users accessing those machines. System queries also play an important role in environments running Oracle software. Oracle’s License Management Services (LMS) auditors have identified a number of “verified” **third-party tools** capable of gathering information useful for managing Oracle license compliance. However, the most common way to gather Oracle usage data (and the way preferred by Oracle LMS) is via a set of proprietary scripts and queries that are made available during an Oracle audit. Unfortunately, Oracle LMS does not make the full set of tools publicly available, though Oracle does offer a free Database Option and Management Pack Usage script **here**.

In addition to the primary discovery tool, companies also should identify a secondary source of information regarding the devices running within their networks. The purpose of this kind of secondary source is to provide a means to measure the completeness of the primary inventory. If there are machines identified in the secondary inventory that are not being captured by the primary inventory tool, then a company knows either that the secondary inventory is over-inclusive and needs to be cleaned up or (more critically) that the company needs to continue deploying or configuring the primary tool to capture all of the company’s computers. Given the limited purpose of the secondary inventory, it may consist of reporting out of a variety of different tools or systems and really needs to consist only of a list of machines and (preferably) the operating systems running on those machines. There are a number of Active Directory queries that can be used in order to generate such a list. Other possible sources of information are lists of machines managed by an anti-virus solution or lists of machines identified within procurement or change-management systems.

Finding the right discovery tool for an IT environment can be a time-consuming and even expensive undertaking, but it is an important one. Most companies need to have a software-based tool for gathering the information required to maintain compliance with their license agreements. However, it also is important to keep in mind the fact that no tool is a silver bullet for compliance. All tools must be periodically tested on a regular basis to ensure that they are working properly. Furthermore, while many tools offer license-management functions, those functions face a dynamic obstacle in the form of software publishers’ ever-evolving license metrics and rules. In order to ensure compliance, it always is necessary to reconcile inventory data against the most current iteration of those rules, and that analysis may need to be a manual one.