Visibility to Workflow: How to Discover and See All Your Technology Assets (for Asset Management Professionals and Architects)
Holistic discovery for all your assets

In order to sustain, build, or transform your business, you need to know the assets that drive your business. And if you’re in the business of tracking assets, you need a clearer and bigger picture of all your assets. There is nothing more frustrating than having only a myopic view of assets in pockets of your business. Organizational leaders demand more, and we should be able to deliver more when it comes to knowing what assets we have.

You can not only build a bigger picture, but a better picture of the assets in your enterprise. The question is how to do that? Start with the data you already have, build a strategy around data you need, work with your operations management team to set up the discovery technology, and create workflows that continue to refine your greater picture of your enterprise technology assets.

This white paper provides an overview of the data you need to see all your technology assets such as software assets, hardware assets, assets in the cloud as well as cloud resources that are often considered assets.

Read on and you’ll find an outline for data discovery strategies and why discovery plays such an important role in the overall technology asset picture. We will also define some of the most common discovery methods and how you can leverage their best capabilities. Finally, we will discuss the importance of bringing all of this data into a single, multisource CMDB and the role that a platform for IT plays in giving you an accurate, trusted, and up-to-date view of your assets wherever they are. You will be able to take that bigger and better asset picture and create workflows that provide immediate savings, improved efficiencies, and reduce your overall risk.
Visibility to workflow—a guide to asset discovery

Asset visibility in the enterprise
Everyone in the organization needs visibility into what is important for their role. However, the higher you go in the organization, a wider view of your technology assets becomes more important because that is where decisions are made, and corporate direction is established. Once those decisions are made and the direction is set, then it’s important to take action, which then filters down into and throughout multiple areas of the organization—and this is where workflows are important. Visibility moves across and up. Workflows move down and throughout. The speed at which you increase the flow of either makes a big difference in the outcomes and livelihood of your organization.

When it comes to their technology asset data, organizations are seeking greater and faster outcomes, many of which include:

- Know what your own, license, or lease, its use, and state of service
- Overall confidence in accuracy of their data
- Real-time or near real-time representations of their environments
- Insights that guide business, set priorities, clarify the next steps, and predict what’s coming
- Quick responses and actions with error-free operations
- Enhanced workflows that refine customer and employee experiences and business outcomes

Understanding the bigger picture…or is it the better picture?
If you’ve ever worked with a color printer, you know that you have various colors of ink. In the printing world they use cyan, magenta, yellow, and black (CMYK) to produce a full color picture. The same inks are now found in nearly every color inkjet printer.

Using the print analogy, how clear, comprehensive, and colorful is your data. From the data you have in the following picture, what is it? And if you choose to Google it, then you’re relying on outside data to answer the question, correct? So, do you know it with your own data?
Seeing a snapshot of only one area of your business is analogous to how people are managing silos of their business or assets. They only see the part of the picture they’re focus on, but it’s only a portion of the data needed to see the bigger picture and answer the important questions. It’s not very colorful either, simply the essentials.

Okay, here’s the bigger picture. Now you have a better idea of what it is, and you may even know where it is. Big city, lots of lights, crowds of people, and pre-covid. It may be a bigger picture, but it isn’t necessarily a better picture, is it? Think of this as having siloed data in one area of your business. You have the bigger picture, but don’t know how it fits together with other areas of the business.

If you want to see the bigger picture with more color, here are the CMYK colors that make up the color picture. Is this a better picture?

If you have all the data you need in multiple systems, you have the big picture, right? Even though we see the picture with all the data, it’s hard to put the picture together in our heads of what the full color picture looks like. Let’s say each one of these colors represents a different set of asset data. We could have cyan representing cloud asset and resource data, magenta representing hardware, yellow could represent your software, and black might represent all the other outside data that gives more definition. It’s all there—but we can’t see the bigger and better picture yet.
How long would it take you to assemble all these colors into one accurate picture? In the printing world it used to take many steps and you had to align them perfectly or they created strange and inaccurate “moire” patterns or a fuzzy effect in the final printed picture (as shown below). You can make out what it is, but you don’t want to look at it long. This is also why you usually have to do calibration tests with an inkjet printer during their initial setup. You could compare this to trying to create your asset picture from data in different formats or normalizing and aligning the data on a similar data model.

With the advent of digital presses and the ability to look at these pictures on your screen, it takes very little time at all to produce the full color picture. Literally and figuratively, these advancements in printing and color rendering are a testament to the power of workflows. And that is what we want to do with our visibility into technology assets—let’s workflow the asset discovery process.

And when we see the entire picture in accurate color, it has greater impact. It’s one of the most famous street intersections—Time Square in New York City.

Data that drives an accurate view of the past, present, and future

In the world of managing technology assets, gathering the full color picture is much the same. We gather data from the past and what already exists, we pull data from current activities with assets, and we try to predict what will happen in the future as we analyze past and current patterns. The asset discovery process, more than anything else, helps us understand what is in our environment and helps us discover patterns that help us predict the future.

Discovery doesn’t help us look into the current operations as much as data that is generated by asset management systems such as software asset management (SAM) or hardware asset management (HAM). However, you put all of these systems together, and just like our color pictures, we can assemble
the right picture and take action on it. When looking at these different data sources here are a few of the questions we are seeking to answer:

- From the past
  - What is currently in the environment?
  - What existing asset data can we bring in to tell us a deeper story about what’s in our environment and how it’s being used?
- Within present operations
  - What actions are happening daily that alter, add to, or remove asset data?
  - What systems or workflows are at work that change or add to our asset data?
- Predicting and preparing for the future
  - What patterns can foretell improvements when we add to or subtract from our environment?
  - How can we create more intelligent patterns to forecast needs and prevent problems from arising?

Visibility starts from data, but where does the data come from? Let’s take some of the types of data related to technology assets and map them to how we can create the bigger and better picture. Four types of data sources include:

1. Existing data within your organization and processes
2. Ongoing data you create because of your processes
3. Outside data that is useful to clarify your insights and processes
4. Data from insights you generate when combining data — in other words, the bigger or better picture, which acts as an input for future decisions and processes

Multiple ways to view your assets

Mention visibility and immediately it conjures up thoughts of dashboards and analytics. But visibility is contextual to a time, a role, or a process. The way that we look at and work with data continues to evolve. With the advent of low code development platforms, such as ServiceNow, there are so many ways to customize your views and interactions with your data. However, there are a few areas of visibility worth mentioning that come out of the box that help you immediately see your asset data, such as dashboards, workbenches, analytics, alerts, service mappings, and workflows.

Dashboards
What dashboards do you have or would like to have? The overview dashboard allows you to customize which metrics are most important to see for you or for any of your roles. Take components for other dashboards and prioritize them on your overview dashboard.

There are dashboards that will help you see what software or cloud resources are costing you. There are dashboards that are included with publisher packs or predefined rules, metrics, and workflows related to each software publisher, such as Microsoft, IBM, Adobe, VMware, SAP, Oracle, and other major publishers. And there are dashboards related to hardware lifecycles that help you see assets and where they are in their respective lifecycles.
There are also dashboards to help you see the discovery process in action with your discovery schedules, devices, software, and cloud resources that have been discovered.

**Workbenches**
Dashboards are nice to keep an eye on your asset estate, but where you can both see and interact with your data are the workbenches. One that many SAM managers spend plenty of time in is a licensing workbench where they can forecast the needs of your organization by trending the number of rights consumed against the number of active rights purchased. You also have visibility into compliance issues such as unlicensed installs and subscriptions or into optimization opportunities such as downgrade and upgrades rights, plus removal candidates. All of which you can take initiate remediation workflows.

**Analytics**
Analytics plays a big part in decision making because of the ability to model usage, trends, and slice data into multiple views. ServiceNow customers use the Analytics Hub as a detailed, exploratory view of indicators. It lets you see trends, predictions, breakdowns, and associated records for a specific indicator.

In the Analytics Hub, you analyze indicator scores by aggregating data, comparing scores, or viewing changes over time, and filter scores by breakdown. You can also enhance your views by adding targets, thresholds, trendlines, and useful comments for significant changes.

**Service mapping**
The complexity of infrastructure and systems paired with the need to maintain critical business services makes service mapping an imperative to understanding your environment and mitigating risks. If one thing goes down, the service goes down and the business or organization loses time and money.

Technology asset data underpins these primary systems and services. The data about the software, the virtual machines running that software, and the hardware that the virtual machines run on. It’s a stack that you don’t want to fall, and service mapping keeps you informed and helps you preempt serious incidents and react quickly to those that come up.

**Alerts**
Alerts can be set on nearly any data point. You’ll want to set alerts on data points you want to monitor that fluctuate where you want to watch certain parameters, such as number of licenses purchased vs licenses in use. You’ll also want to set up alerts that help you keep on top of things like your hardware lifecycle. This would include approaching end of life or end of service within three months, software contracts nearing renewal or expiration, SaaS subscriptions, and leased hardware coming due. All of these factors pose risk, but other alerts are set primarily to sniff out risks such as those related to potential hardware failure, unsupported or vulnerable software versions, or compliance alerts.

Alerts should be easy to customize and interact or leverage such things as IT workflows, business rules, scheduled jobs, and separate notification systems.

**Workflows**
Business priorities for better experience, efficiencies, and innovation drives the creation of workflows; but it is the data that daily powers those workflows. Asset data is needed in workflows such as HR onboarding and off-boarding, asset
lifecycle management, ITSM incident and response, and security vulnerability and remediation. Each of these workflows uses technology asset data that comes from either discovery or data generated from asset management solutions. The more accurate and timelier the data, the better the experience, savings, and ability to innovate.

**Discovery strategies see deeper into the past**

**Why discovery**

In the broader sense of data visibility, you could look at discovery as the black layer of our picture. Part of the discovery process is bringing in data that is discovered with other tools to fill in more detail, similar to the difference between a black and white bitted graphic and a gray scale photograph. The details make a difference. Likewise, with a good discovery practice, we have a clearer idea for what’s happened in the past through discovery on multiple fronts.

A large portion of our overall picture comes from discovery. Discovery enables the visibility of almost all areas of IT including, security, operations, asset management, and service management. But there are so many facets to discovery and making sure the data is clean, clear, and current. Let’s break down the discovery processes for enterprise visibility and look to see how it empowers IT and asset management.

**Strategy for discovery**

The discovery strategy we put together should cover not only what we need, but our approach to getting the right data without causing more confusion. Your discovery strategy for asset management should include the following four areas (note—the first two are about the what and the second two are about the how):

1. Prioritize your data—establish which data is most important to your organization for:
   a. IT and business insights and decisions
   b. Processes and to maintain business continuity
   c. Security
   d. Compliance

2. Define your data authority— in order to establish which sources provide what data, you’ll need to answer questions related to which:
   a. Data do you prefer, which is tied to the first step of your strategy
   b. Data do you trust
   c. Data source has the authority
   d. Fields in your configuration management data base (CMDB) need to be updated by what data

3. Determine your data structure, storage, and access—much of this has to do with the setup, maintenance, security, and state of your CMDB.

4. Utilize data layering from multiple sources for deeper dives—which is especially important for organizations that are managing and using data that goes deeper—technology asset management/ITAM, or security operations, or business operations.
Executing your comprehensive asset discovery projects

Discovery is usually the domain of IT operations management (ITOM). However, IT asset management project (ITAM) and ITOM discovery projects often go hand-in-hand as ITAM heavily relies on discovery tools to collect all relevant, asset-related data. An ITOM discovery project is a project on its own and requires a specific skill set. An ITAM project has specific requirements for discovery, such as Oracle database discovery or integrating discovered data from an endpoint management system such as SCCM or Jamf. However, it should not be assumed that the ITAM project takes ownership of the discovery process, unless specifically agreed upon with the ITOM team.

Instead of mixing the two projects, it’s better to align them. The ITAM project becomes a stakeholder of the discovery project with its own set of requirements that should be accounted for within the discovery project plan.

Unifying data from multiple discovery methods

There are several methods for discovery, and most were built with specific purposes in mind. Here are a few of the more popular discovery methods:

Agentless

If you need to know what goes on within a datacenter, to monitor operations, or to build mappings of dependencies and systems without being intrusive, agentless was developed for this purpose. The ability to discover is mostly accomplished through protocols and queries used on the internet with light code that sits on a server on the network. Agentless discovery is very good at discovering devices that don’t move around as much and the relationships between devices, systems, and processes.

Agent-based

Designed to do deep discovery on endpoint devices, which requires a certain level of code on the end device being discovered. This discovery method often comes as an adjunct ability added to agents most commonly used to manage systems and configurations on devices. These agents are built by management software manufacturers to manage a variety of device-related activities and provides not only deeper data, but also gives you greater abilities to take multiple actions on that device. Often useful to discover devices that move around the network and pop in and out of the network.

Unified endpoint management – is similar to agent-based, but usually has lesser code, because it leverages APIs to management capabilities that have been exposed by each OS vendor. So, instead of having a larger client or agent on each device that does all the work, much of that work is shifted to management actions native to the OS.

Sensors, collectors, or listeners

These discovery methods are similar to the agentless discovery but are developed to watch for certain activities on the network or on specific devices. They often use a combination of internet protocols to either query a device in real time or to site on sub-nets around your network. Like cookies on your browser, think of them as the cookies or maybe cookie crumbs that lay about you network gathering information. They can be on server-like devices or on all the devices on the network, and they require very little code.

Unified Agent Strategy

ServiceNow’s vision for agents is to provide an agent client collector that leverages the power of the Now Platform to aggregate, normalize, and synchronize data across all areas of IT and the business experience. This means one agent can be used to gain visibility and trigger workflows within service management, security operations, event management, and asset management.

This unified agent strategy will reduce the need for many other agents to gather discovery data needed within the ServiceNow platform. However, it does not replace agents used in workflows to reclaim or remove software from endpoints, as is the case with SCCM or Jamf.
Event-based triggered discovery

Use identifiers to compare configuration items (CIs) within the CMDB for CIs with matched discovered devices. These identifiers can be configured to detect changes. When IT uses VMware, AWS, or Azure to make changes to their virtual, or cloud environments, these events trigger discovery to update the CIs and corresponding relationships. This ensures up-to-date accuracy of the CMDB and real-time visibility into virtual and cloud environments.

- Pattern-based discovery such as file-based inventory or customized patterns that find gaps other general discovery methods don’t
- Cloud or hybrid

Connectors between discovery methods

As an example of what can be discovered, this table represents some, but not all, of the discoverable infrastructure through a combination of discovery methods:

<table>
<thead>
<tr>
<th>Discovered infrastructure</th>
<th>Web and database servers:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computers, clusters, and virtual machines:</td>
<td>Apache web servers</td>
</tr>
<tr>
<td>• AIX computers</td>
<td>• General software packages</td>
</tr>
<tr>
<td>• ESX servers</td>
<td>• HBase on UNIX</td>
</tr>
<tr>
<td>• HPUX computers</td>
<td>• JBoss servers</td>
</tr>
<tr>
<td>• Linux computers</td>
<td>• Microsoft IIS servers</td>
</tr>
<tr>
<td>• Linux kernel-based virtual machines (KVM)</td>
<td>• Microsoft SQL servers</td>
</tr>
<tr>
<td>• Mac computers</td>
<td>• MongoDB servers</td>
</tr>
<tr>
<td>• NetWare computers</td>
<td>• MySQL servers</td>
</tr>
<tr>
<td>• Solaris computers</td>
<td>• NGINX Web servers</td>
</tr>
<tr>
<td>• Solaris zones</td>
<td>• Oracle database</td>
</tr>
<tr>
<td>• Virtual machines</td>
<td>• PostgreSQL instances</td>
</tr>
<tr>
<td>• VMware vCenter</td>
<td>• Tomcat servers</td>
</tr>
<tr>
<td>• Windows computers</td>
<td>• WebLogic application servers</td>
</tr>
<tr>
<td>Hardware devices:</td>
<td>• WebSphere servers</td>
</tr>
<tr>
<td>• Dell DRAC</td>
<td>• Other technologies:</td>
</tr>
<tr>
<td>• Load balancer’s and load balancer proxy software</td>
<td>• Amazon Web services (AWS) cloud</td>
</tr>
<tr>
<td>• Network printers</td>
<td>• Connections</td>
</tr>
<tr>
<td>• Routers</td>
<td>• IP networks</td>
</tr>
<tr>
<td>• Storage devices</td>
<td>• Puppet automation software</td>
</tr>
<tr>
<td>• Switches</td>
<td>• Relationships</td>
</tr>
<tr>
<td>• Uninterruptible power supplies (UPS)</td>
<td>• Services/Daemons</td>
</tr>
</tbody>
</table>

https://www.youtube.com/watch?v=tPfIVTsJBGg 13:46
Better picture from a complete asset discovery

We’ll look at why it’s important to do discovery in three different areas to build a cohesive technology asset picture from your software, cloud, and hardware estates. Then we’ll explore what kind of visibility we need to make decisions and understand our environment. And a fair amount of our discussion will encompass a look at the main discovery technologies used in each area and some that are under-utilized because they’re new or lesser known.

Software asset management discovery

Recommend a holistic approach

There isn’t one discovery method that catches all software in the enterprise. You need to take a holistic approach to gain full coverage while discovering software products. One discovery method often does not work in every scenario. For example, if you are running a process-based discovery you have the risk of not discovering software when the process is not running or does not have a registry entry. Your software discovery should watch for not only executables to run but check file directories to match with known software file names, and have pre-configured patterns in place for large, well-known software vendor software with a library that is continually updated.

Why trusted software data matters

It’s important to find up-to-date configuration items, assets, and file-based patterns on-premises and throughout your cloud infrastructure in order to mitigate risk and reduce costs within your organization. Discovered software data helps reduce your software license risk by helping embed software asset management into IT change and being able to act on unlicensed or seldom used deployments.

The same ability to scan for on-premises software should extend to cloud software to cut cloud spend by finding shadow IT, reduce purchasing overlap from rogue purchases, and optimizing licenses use across the enterprise.

Software data also feeds into systems that help identify vulnerable software with information from the National Vulnerability Database (NVD) and can help you remediate vulnerabilities with automated workflows for Security Operations.

Applying discovered software data with workflow lifecycle dates from the content library identify or alert you to software end of support that further put you at risk. Other alerts generated from discovered data that meets thresholds can also show you which business applications have demand risks that need to be addressed.

Common methods for discovering software

Discovery of installed software is at the core of an effective SAM program. ‘Completeness in discovery’ and ‘accuracy in software recognition’ are key factors in this process. If the discovery tool is unable to find a software installed on a server or workstation, it will result in an in-accurate true up cost calculation and risks during audits.

One effective method of software discovery is scanning or probing across the network for all software that executes on a device or the network. This usually happens with scheduled scans that occur across certain IP ranges, the software is classified, then identified with objects in the CMDB, and probed further for more information. (See sidebar for a ServiceNow Discovery example.)
There are times within controlled environments where running advanced queries or connections using certain protocols are restricted. To work through all situations your discovery must leverage multiple methods and technologies to ensure complete coverage. Certain software doesn’t even require installation; all someone has to do is deploy the executable and run it directly from command line or terminal. A standard software scan will most likely not catch these use cases.

**File-based discovery catches what other methods won’t**

File-based discovery finds files on Unix, Solaris, and Windows machines and then matches the file name and other discovered file properties (such as file size, etc.) against a set of pre-defined normalization rules to correctly recognize the software publisher, product, and version. Content libraries that are built for SAM purposes are used to compare the file-based patterns for different operating systems. You can point it at the OS file structure and then tell it what folders to ignore and which should be blacklisted.

The file-based discovery can find things that might not be found in a regular package manager scan. Items that are not able to normalize initially are stored and contain file signature information including file name, and file size. Once these are manually normalized, they will transfer the publisher, product, version, edition, platform, and language. They can be manually normalized when the SAM content library hasn’t identified it yet. File-based discovery also helps update the content library without writing new patterns to help with future discovery and normalization.

File-based discovery fills in the gaps where software package discovery comes up short, such as when software is loaded on Unix and Linux but isn’t registered or isn’t going through the standard install methods. This type of software is hard to detect. Other difficult patterns to detect, which file-based discovery can catch include code that runs within the package, but isn’t the main software, some of these ancillary processes can lead to security issues.

If file-based discovery can’t classify a particular software, it will represent the software in a bucket called general software for you to assign or normalize. Once this is done, those new patterns become part of the classifications for future scans of that software.

Like other methods, file-based discovery has its shortcomings. It can be resource intensive, but it is worth running after a software package scan to fill in gaps for specialized situations. It will also provide additional patterns for future scans, as well as add to a comprehensive SAM content library.

**Caution:** Be selective at which folder structures you point file-based discovery at to scan, because if it is pointed at all file systems or very large ones, it can tax your network and SAM system and crash your CMDB.
If you would like to take a deeper look into the discovery pattern from ServiceNow see the documentation site under ITOM visibility that lists all the types of data collected under “available discovery patterns”.

![Diagram of file-based discovery](image)

**Figure 2: File-based discovery**

When you use file-based discovery, you don’t have to build patterns as much and the patterns it does discover will continue to enrich the content library for future discovery and discovery through other methods.

**Advanced patterns for software asset management**

Software publisher packs—Discovery helps create many of the patterns that show up in many of the publisher packs. In the case of ServiceNow, a majority of the publisher packs leverage discovery data for compliance and optimization. These out-of-the-box publisher packs include constantly updated patterns, dashboards, and compliance reports and the ones that leverage discovery include:

- VMware (driven almost entirely by discovery)
- Microsoft
- Oracle
- Citrix

Customized discovery patterns—In order to customize the data collected, you need an advanced pattern designer. This allows for quick updates to collect new attribute data not already being collected out-of-the-box. But it also allows for quick massaging of the data and insertion into the CMDB. It can also leverage data that already exists in the CMDB tables.

Machine Learning and AI—Discovery plays a large part in providing data for machine learning and AI. Without it, there would be very little to initially run patterns and algorithms against. Machine learning and AI are in many ways the next level of discovery—watching the past patterns to predict future behaviors and results.

**Cloud applications and resources discovery**

In many cases, organizations don’t have a good idea of what they have in the cloud, so first and foremost, you need to gain an awareness of what has been
provisioned, the accounts you have, and what is out there. This not so simple step is the foundation for everything you’ll do to manage assets and resources in the cloud and the awareness alone will inherently reduce shadow IT and cloud spend.

**Value of cloud discovery continues to grow**
Visibility into your cloud infrastructure is critical for managing your operations, optimizing your cloud spend, security, bring your own licensing, and compliance. Cloud discovery also provides visibility to these resources for all the major cloud providers such as AWS, Azure, Google Cloud, VMware Cloud and IBM Cloud.

Once you have a consolidated view of your cloud footprint with things like SaaS, IaaS, and PaaS consumption, then you can implement operational best practices, compliance, security, and cloud spend optimization. Until you’re keeping an eye on your cloud activity, waste and vulnerabilities will continue to run rampant.

**Two popular cloud discovery methods**
There are two main types of cloud discovery. First, there is service account discovery which finds resources by collecting meta-data using the cloud providers APIs. These APIs push information in response to cloud events, so this kind of discovery keeps your CMDB updated for short lived or ephemeral cloud objects in new real-time.

The second method is IP-based discovery, which enriches the metadata from service account discovery with deeper data like the file system, installed software, and process information. IP-based discovery also detects enterprise applications like Oracle and SQL databases and Apache. This data supports a variety of IT workflows.

Discovery also uses cloud tags to add metadata on the application, cost center, and environment. And service mapping processes use these tags to automatically create application service maps. Cloud resource discovery supports AWS organization, Azure management groups, and Google resource hierarchy, which helps you identify master-to-member account relationships.

**Keeping cloud asset data fresh**
Because the majority of cloud workloads are short-lived, it is imperative to have the CMDB reflect the near real-time nature of cloud workloads to effectively automate workflows. Your discovery workflows need to support event-driven discovery where events from the cloud provider keeps the CMDB up to date. This type of discovery is best done through event-driven discovery. Customers can push cloud events that trigger discovery on demand to the target service, which in turn identifies changes and updates the CMDB in near real time.

It’s also important to see, set, and add discovery schedules. You should be able to focus discovery schedules by region or by scope, such as whether or not you want to discover all VMs or have it do it through other means, such as SSH and PowerShell. And of course, you want to run these discovery schedules on off hours in each region, such as in the early morning hours.

Use dashboards that show current discovery schedules, discovered devices, cloud service accounts, and any discovery errors. View cloud resources shows the cloud resources for all your providers and more details sorted by provider.

Metadata discovery often helps customers with basic data visibility on VMs. However, the data is not usable for outcomes like software asset management or AIOps or SecOps use cases. The compute layer should be interrogated by IP-
based discovery schedules to collect enriched data sets, such as installed software data, process information, TCP/IP connections, enterprise applications deployed on cloud VMs, and so on.

**Bring your own license (BYOL) help minimize transitional costs to the cloud**

Organizations need to manage the licenses brought to the cloud. However, there are not many automated mechanisms to determine the license types of resources purchased directly from the cloud vs. brought to the cloud. Instead, people often use “manual” tags on the cloud to segregate these license types. The manual tags added by organizations are inefficient mechanisms for managing this transferrable licensing.

Bring your own license is a way to discover software in the cloud and align it with transferrable licensing allowed by the software vendors. It is better to automatically discover these types of licensing with vendors such as Microsoft products—Windows Server and SQL Server on AWS and Azure. Microsoft licensees can be automatically discovered without any reliance on manual tags, through a combination of APIs and billing information from various cloud providers.

License types across cloud resources can be automatically determined with automated tags added to discovered virtual machines. These automated tags signify whether the license type is either license included or BYOL.

There are certain cases where the cloud providers don’t offer any APIs to determine the license type. In these cases, solutions can infer the license type automatically through billing data downloaded from the cloud provider.

**Multi-cloud and hybrid**

For many customers, migration to cloud is a transformational journey. Visibility to multi-cloud/hybrid cloud deployment data with on-premises infrastructure and application data is critical to solve the real-world IT operations challenges. The CMDB continues to evolve as the single source of truth for any data requirements related to ITOps or CloudOps use cases. The discovery solutions need to offer near real-time visibility to multi-cloud IaaS/PaaS/ PaaS services across private, public, and hybrid clouds.

**Insights to cloud resources and active sessions**

Agentless solution—the cloud discovery that you’re looking for will help you discover data related to the infrastructure and what is running at any given time.
Installed software discovery - Get visibility to packages deployed in servers. Package rules from the SAM normalization engine are used to normalize the content. Discovery solutions help to detect product name, publisher name, version, and edition information. Other cloud information that can be discovered includes:

- CPU information—visibility to number of cores per CPU / number of CPUs
- Visibility to virtualization—see the software installations in virtualization technology like VMWare/HyperV
- Visibility to clustering—software deployed in clusters like Windows, Oracle clusterware
- Visibility to cloud—software deployed in AWS/Azure with the topology
- Tag discovery

Service account-based discovery collects tags associated with IaaS/PaaS components in CMDB key value pair tables. Tags are labels in the form of key-value pairs that may be attached to cloud resources, such as instances, storage volumes, and databases. Tags provide additional information and context about a specific resource.

Real-time visibility into tag data in the CMDB helps with workflow automation and data reporting requirements use cases.

**Container discovery**

Enterprises are turning to containers to increase efficiency, facilitate application portability, and enable multi-cloud IT operations. Enterprise adoption of containers and container orchestration products like Kubernetes, OpenShift, Cloud Foundry continues to grow rapidly in 2019, transforming how applications are created, deployed, and managed.

These containerized applications can be delivered to cloud, on-premises, and hybrid environments. Kubernetes has now become the de facto standard for deploying containerized applications at scale in private, public, and hybrid cloud environments. It’s because the largest public cloud platforms AWS, Google Cloud, Azure Cloud, IBM Cloud, VMware Cloud and Oracle Cloud now provide managed services for containers that are focused on container discovery within the context of cloud discovery.

![Figure 4: Kubernetes flow](image-url)
Visibility to modern technology stacks drives quantifiable outcomes to reduce cost and improve productivity for the CloudOps/DevOps and SRE teams. Container discovery of Kubernetes stack provides near real-time visibility to all the Kubernetes taxonomy data in the CMDB.

Event-driven discovery provides near real-time visibility to Kubernetes data in CMDB. Discovery pulls the event data based on the last execution timestamp. The frequency of event-driven discovery is configurable and hence customers with ephemeral workloads can collect near real-time insights in CMDB.

Red Hat OpenShift is one of the more popular Kubernetes platforms and container types on the market today. OpenShift provides all the capabilities for IT operations teams to deploy and manage containerized applications including the Kubernetes distribution itself, but also the container runtime, storage, and networking capabilities, logging and management, authentication and security capabilities.

Labels are key/value pairs that are attached to Kubernetes objects, such as pods. Labels are intended to be used to specify identifying attributes of objects that are meaningful and relevant to users, but do not directly imply semantics to the core system. Discovery detects label data and populates these elements into the CMDB key value pair table.

Container platforms are a hot topic in the enterprise. They represent a top investment theme in the executive stories of digital transformation. Cloud Foundry continues to have a dedicated loyal group of enterprises using their platforms to run PaaS containers.

Discovery needs to support the discovery of cloud foundry cluster, organization, quota, route, space, services, applications, and other critical elements from the PCF layer.

Hardware asset management discovery

Outside of the frustration or embarrassment of not knowing what you have or where things are, hardware asset management (HAM) simply reduces costs through better planning, procurement, and contract negotiations.

A large part of that cost reduction comes through automating your hardware lifecycles and knowing where assets are within each stage of the lifecycle. Another important aspect of HAM is the ability to minimize risks due to asset loss, security vulnerability, compliance, and environmental accountability and concerns.

Clearer decisions with cleaner, more reliable and trusted hardware data make any IT department look good and act even better.
Hardware from four vantage points
When looking at hardware asset management, it’s good practice to view it from four different vantage points while planning and evaluating your discovery needs. These vantage points include the following:

- Endpoint computing
- Network computing
- Server computing
- Virtual infrastructure
- Combination approach

Each of these has pros and cons for whether to prioritize agentless or agent-based discovery. Hardware that has static IP addresses, such as network devices or hardware that doesn’t move between network segments often—these are better suited for agentless discovery.

Hardware that moves around such as laptops or mobile devices are better found by agent-based discovery, because the agent is on the device and reports in when it attaches to the network on regular intervals vs. being scanned on set schedules.

![Figure 6: Asset normalization engine](image)

**Endpoint computing**
Desktop computers that don’t move around work well with agentless discovery, but their usage overall is diminishing. However, with more people working from home, even laptops are staying put more often.

In an office environment, laptops are more difficult to get consistent data when scanned agentlessly on a scheduled time frame. The mobility of laptops is appropriate for persons that have to move between locations and offices throughout the day. As the laptop moves, the subnet and IP address to which the laptop connects changes as well. This continual change of IPs makes consistent, agentless discovery difficult. Otherwise, it’s easier to capture information from laptops from agent-based discovery, which can be done by an
agent or through connectors to various endpoint management agents. Another option is to capture information from laptops via agentless discovery. Laptops during off-business hours usually sit idle on docking stations or in one location during an extended period. This makes agentless discovery effective during these off hours.

User—both assigned to and primary user—information is often best done through an agent, which can detect who is logging in and out of the device.

Network computing
When it comes to network computing, agentless stands out more because you can capture a lot of data about the networking relationships and be able to build a service map between the networking devices and the applications that run on or are dependent upon those devices.

Many agents don’t collect this type of relationship content unless the agent was specifically designed for a particular brand of network gear, such as Cisco. Combining data from various networking device scanners can be beneficial for a complete inventory of network devices but will be inaccurate when trying to combine relationship information, since each scanner will capture network relationships with different methods.

Server computing
Like network computing, servers don’t move around the network, so agentless discovery works better; plus, many people in the data center don’t like a larger agent on all their servers which, if the agent malfunctions, could potentially bring down a server or a service.

As part of that service mapping, the servers as well as the services and applications can be mapped with the data coming from agentless discovery. However, there are agents that are designed and verified for datacenters and connectors to those agents can provide a wealth of additional information, such as the sub-capacity usage and software metering data.

Virtual infrastructure
Because virtual infrastructure changes quickly, continual discovery of your virtual devices is critical to managing licensing and risk in your environment. Virtual devices can be discovered directly via the device’s IP address or through the virtual management servers that centrally control your virtual devices.

Pulling in virtual devices outside of discovery through connected tools will bring an inventory of virtual devices from those specific tools into your CMDB, but it will not help with the most important part of your virtual infrastructure—mapping the use your virtual devices provide to your critical business services and applications.

Discovery not only creates the relationships between the virtual devices and their host servers, it lays the foundation for service mapping—mapping the use your virtual devices provide to your critical business services and applications. With this complete and accurate service map, you are able to determine lifecycle implications on your critical services and applications. The software and hardware can be properly maintained, patched, and replaced in accordance with minimizing risk.

Combination approach
In most cases, we recommend a combination of agentless and agent-based discovery with a data prioritization or data bifurcation—where one discovery...
method is used to update certain fields and another discovery method is used to update other CMDB fields—that is set up in your discovery practices.

Examples of why you want both methods include when an agent isn’t up to date or not reporting in for some reason, agentless discovery can still provide a certain level of data. Another example would be if a device is moving around a lot and isn’t being discovered by an agentless scan, the agent can make sure that device is accounted for, plus it can provide more in-depth data about software usage.

Hardware discovery goes beyond discovering devices, it can be helpful at providing data at the component level. For example, you may want data from a security standpoint to track the serial numbers for each hard drive in your computers to correlate the hard drive to the computing device. This will help you see if any hard drives are missing or have been swapped out or switched to another device or laptops. This ability to track storage devices is especially necessary in highly regulated or secure environments.

Bringing in 3rd party discovery data

The reason why end user devices are so difficult to track and keep updated with agentless discovery are the following:

- Devices are not always on and available during a scheduled discovery run. The end user may turn off the device or not be connected to a network where it is accessible remotely by agentless discovery.
- Devices are often locked down and made inaccessible for remote access. Good security policy should protect end user devices, but if the devices are not accessible, then agentless discovery cannot connect to that device.
- Security policies change, such that devices that were once accessible are no longer accessible, or credentials that worked previously are no longer granted to you.
- When the CMDB doesn’t have the latest data from a particular discovery source and you choose to reprioritize data from other sources.
- The mobility of end user devices produces a complexity in itself. Devices that move throughout the day from one network segment to another can evade discovery by the nature of constantly changing IP addresses.

It’s because of these reasons and the additional value and depth of data connectors can provide that we think it is important to include this data in your overall asset discovery processes. What follows is a short discussion of other discovery sources for technology asset data.

Unified Endpoint

We can benefit from more discovery data through the evolution of endpoint management to unified endpoint management. There is an evolution and consolidation happening for capabilities that manage endpoints like laptops, desktops, servers, mobile devices, and now even edge and IoT devices.

The same thing is true with discovery technologies, which continue to advance, but are tightly tied to the type of endpoint being managed. These management tools have as much information about the endpoint as they are provided through OS management APIs. That’s why it’s important to evaluate them for the type of data they can import into the CMDB.
Many of the endpoint agents we’ll discuss are now using unified endpoint management technologies and APIs. There are lighter management, agent-based discovery tools that have been developed completely with unified endpoint management in mind, especially when it comes to mobile computing such as smartphones and tablets. Intune is an example of this type of light management agent that collects hardware and software data. Using the data from an Intune connector, we can gather:

- Detailed hardware and app inventory tracking for Android/Apple/Windows mobile devices
- Software on device
- Compliance tracking for mobile devices
- BYOD or corporate owned devices

**Endpoint or client management agents**

Endpoint or client management agents often sit on laptops and desktops, and sometimes on servers. We will highlight a few of the agents and what they discover below, but this is by no means a comprehensive list, rather only a few of the major player.

**Microsoft’s System Center Configuration Manager (SCCM)**—is a common agent that sits on Windows devices. When you use Service Graph connectors to SCCM, you can gather data such as:

- Detailed hardware and software inventory tracking
- Users
- Datacenter infrastructure and server management
- Ability to detect delta changes for efficient incremental imports from SCCM
- Detecting removed software
- SCCM’s Asset Intelligence

The connector along with the SCCM agent supports ITAM workflows delivering data used for greater visibility, compliance, governance, purchasing, and planning. A few of the workflows it enables include:

- Inventory for software package and installation tracking
- Compliance tracking
- Software usage tracking
- Software editions for normalizing publisher/product information
- Remove reclamation candidate software from the device
- Integration with SAM and client software distribution workflows

**Jamf**—is a widely used agent in the world of Apple devices such as Macs, iPads, and iPhones. Other companies like Microsoft and IBM have teamed up with Jamf to fill out their unified endpoint management practices. Using a Service Graph connector, we can gather the following endpoint discovered data:

- Hardware and software inventory data
- Software package and installation tracking
- Detailed Mac hardware and software inventory tracking
- Detailed iPhone and iPad hardware and software inventory tracking

**ServiceNow Jamf integration**

Import hardware and software data from Jamf Pro according to CMDB and ServiceNow SAM best practices and enable your help desk and asset managers with:

- **Additional data**
  - Uses same tables as other CMDB/asset integrations
  - Imports and merges with Jamf data for publisher names missing from the Jamf API

- **Import policies**
  - Import Jamf policies to use with pre-built workflows for request, approve, fulfillment, and removal

- **Easy setup**
  - Guided setup, import wizard and dashboards
  - Bring in Mac and device data (iOS and tvOS)
  - Out-of-the-box tables for (cmdb_ci_computer, cmdb_ci_comm, cmdb_sam_sw_install), which doesn’t add lots of CI columns
Much like the SCCM data, the Jamf data helps provide the answers to many ITAM workflows, including:

- Tracking software package and installations
- Creating software catalog items with pre-built workflows that automate the request, approval, and fulfillment of software from your service portal
- Tracking compliance for mobile devices and end user computers
- Automating installation and removal for onboarding/offboarding or software requests
- Supplying software usage data to reclaim underutilized software
- Automating the removal of under-utilized software
- Delta change detection to reduce the daily import and performance improvements
- Managing Macs with the same processes and level of detail as Windows or other endpoint devices

**Other notable endpoint discovery solutions**—There are other agents or light agents that help discover data. Two worth noting are the BigFix agent and Tanium. The BigFix agent brings in detailed hardware and software inventory tracking for Windows and Linux hardware and software. Tanium uses different ways to query devices on the network in real time, and by doing so, Tanium can bring the following to your overall visibility and asset data story:

- Discover data about workstations, servers, and cloud instances, whether that endpoint is on or off the corporate network.
- Pull external data from 3rd party systems into the CMDB (e.g. warranty information).
- Provide real-time data from all endpoints, without needing additional agents, supporting systems, or administrative credentials.
- Create scripts that flexibly query endpoints at scale, which is helpful to find non-traditionally installed applications.

BigFix and Tanium and the other Service Graph connectors can bring asset data into the CMDB with or without an ITAM or SAM solution; however, the majority of pre-built workflows related to asset management are found in the ITAM and SAM solutions. The data from these and other endpoint agent-based discovery tools can be integrated into many of the same workflows previously mentioned.

**IoT tools**

Armis and ForeScout are two of a growing number of solutions that discover IoT and non-managed devices. Each is focused on making sure these devices don’t pose a threat to your network and they provide discovery information about the IoT devices. When needed, Service Graph connectors bring this data into the ServiceNow platform, so the devices can be tracked as assets and managed within IT workflows.

**Cautions and caveats of only using Service Graph connectors**

Unfortunately, using only connectors to populate your system of record could make the job of building a healthy CMDB more difficult. In our “bigger and better picture” analogy, building all your data sources from connectors is like our black and white or grayscale photograph, we know what it is, but it isn’t as vivid
or an accurate enough depiction of our environment. Adding a bit more color to this: in a survey of customers, two of the biggest issues they faced with building a CMDB was stale and duplicate data, which can happen when relying exclusively on third-party data sources.

As part of your overall discovery strategy, it is important to know the scope of discovery from your third-party discovery methods. You need to know what each method can and can’t discover and understand how the discovery data is stored. Knowing the scope will help you determine:

1) If the third-party source will provide you with configuration item (CI) data your primary discovery tool, such as ServiceNow Discovery, doesn’t.

2) If your CMDB will gain value from the third-party discovery data, then knowing the scope of the discovery will help you decide which specific data to bring in and what data not to bring in. You only want to update your CMDB with information that will make your CI data more accurate and complete.

3) Whether the third-party discovery provides more near-time or real-time data from devices or sources that require more frequent data.

**Stale data**—Configuration item data is useful and actionable when the data accurately captures the current environment. This means that CIs stored in the CMDB should be refreshed in a timely manner. When the CI data is not refreshed in a timely manner or the CI does not represent the current environment, the CI data is considered to be stale. When CI data is stale, configuration items may no longer exist in a customer’s environment, even though they are reflected in the CMDB. This is also the case when the current configuration of the item in the environment has changed and those changes are not represented by the CI.

If the data from the third-party source does not have up-to-date information about the resources, then when the data from the third-party source is brought into the CMDB, the CI will be updated with data that is actually days, weeks or months out of date and after the update does not accurately reflect what is in the environment.

When discovery runs according to one or more schedules to collect data, we know it is accurate at that time; consider it a snapshot of your environment at the time of discovery. Customers demand this especially for processes such as operations (e.g. event) and security (e.g. security incident) that require timely and accurate data.

Up-to-date data in your CMDB is so important that you should consider incorporating event-based discovery to keep the CMDB up to date in near real-time. Your CMDB would receive events from public cloud providers and virtualization hypervisors when a change takes place to a resource in the data center and the discovery process is launched immediately to refresh the full data from the affected CIs with the most updated data and network connections. You can also enable CMDB Health to monitor the health of your CMDB. When stale CIs are identified, a remediation task to handle each stale CI is created. Finally, data certification can be employed where stakeholders periodically review the content and quality of the CMDB, and update data as needed.

**Duplicate data**—Customers must ascertain the quality of the data being brought in to their CMDB to avoid creating duplicate CI data, which could also be a
problem when using a connector. This comes from how you apply your data prioritization from your strategy into action:

- While Service Graph connectors are built to use the robust transformation engine (RTE) and the identification and reconciliation engine (IRE) to specifically address the duplication issue. Duplicates in the data are determined where unique values are matched to multiple records. These engines won’t prevent duplicates from being brought in if the third-party source has duplicate data with unique values that cannot be matched to existing data.

- This means a customer should take care to ascertain the quality of the data in the third-party source before bringing in that data even using a Service Graph Connector. This is a best practice that should be followed anytime data is brought in from an external source.

**Limited data**—If a customer is looking to populate an accurate and service-aware CMDB, bringing data in only via Service Graph connectors is not the right approach. It is recommended to have one or two main discovery methods in place to do the bulk of your discovery and then augment them with Service Graph connectors to pull specific data from devices and systems. Without a mainstay discovery process, outside of Service Graph connectors, you will not be able to accurately and completely map your services for a complete enterprise-wide view.

The reason is that certain discovery and visibility tools, such as ServiceNow Discovery and ServiceNow Service Mapping, have been specifically engineered over a significant time to collect data from numerous technology devices and interpret this data to populate the appropriate classes and relationships within a CMDB.

**Connector community**—Having a community of vendors along with dedicated teams to develop Service Graph connectors helps customers leverage a better overall data visibility experience. There are similar setup and operations between all third-party connectors with the ability to use a common services data model (CSDM) to move data to and from the CMDB. The community helps streamline the process for creating connectors for partners (ISVs), which will increase the number of asset-related data points you can use to build the layers of your visibility and the intelligence in your workflows.
Role of the CMDB with asset discovery

It is especially important to have an accurate CMDB to be able to do workflows. If you’re using a method that pulls data from different sensors, agents, or other sources and then tries to apply it to a workflow—that workflow will always be constrained by the slowest connection or the missing data if there is a problem with the connection. Using a CMDB is a great way to make sure that the data is available to provide workflows with accurate data. The bonus that comes from a healthy CMDB is ensuring the data flows in a timely manner, is normalized and reconciliated, and is readily available for consumption by workflows as well as IT management tools and processes.

Accuracy is of utmost importance considering how many critical business processes depend on accurate data: asset management, operations, service management, compliance, security to name a few. Given the scale of this problem and importance to customers, ServiceNow has undertaken a proactive effort to help customers build and maintain an accurate CMDB.

An accurate CMDB requires management, instrumentation, and a discovery (MID) server, which can run on any server in the environment and or to connected resources like cloud infrastructure, apps, and devices. And an advanced CMDB also recognizes relationships between devices, apps, and services, which helps to build dependency mapping to apps and to services.

Effective discovery requires a multisource CMDB

Getting back to our bigger picture analogy, think of a multisource CMDB like Adobe Photoshop layers and how you can add more with the introduction of another layer of discovery data. These additional layers allow for multiple data sources to add deeper data to each configuration item. This is important for those practices that leverage deeper data or broader sets of data in order to make decisions and to take action.

Many of the attributes of a multisource CMDB work exceptionally well to improve discovery results and make sure the date is accurate, validated, and prioritized. Use the multisource report builder to create queries for multisource CMDB records, discovery sources, and CI records and gain further operational insights about multisource CMDB data.

You can also validate a new discovery source by comparing its data to data from other previously validated discovery sources. Feel free to drill down into the discovery sources at the attribute levels or modify reconciliation rules and then recompute CMDB data, reflecting the updated reconciliation rules.

Many might fear doing something wrong with their data coming into the CMDB, but you can revert CMDB data integration from a specific discovery source—for example, if you realize the discovery source is not reliable, you can recompute CI attribute values, while excluding the discovery source you want to ignore.
Near final thoughts about asset visibility workflows

There are no final thoughts when it comes to discovering, monitoring, and managing your assets. It’s constantly changing. And there isn’t a single method or technology for discovering all your assets in your enterprise. The best place to start is with a strategy for defining what data sources you trust, what data you need, and which discovery methods will give you that data. The execution of your strategy will help build, define, and clarify your overall asset picture. There are different methods for discovering data on your network, for devices that move around your network, and for assets and resources in the cloud.

There is, however, a way to bring data from multiple sources and transform it into a single data model and use the discovered data from multiple sources. Plus, there is one place you should store all the information discovered about your assets—a multisource CMDB built on a platform that makes it easy to share data, create workflows, and develop processes and applications that leverage all of your data.

We can keep our heads down and focus only on our own little sphere of assets, or we can broaden our vision, clarify what we have enterprise-wide, and add color to what we can do with our assets and resources. But we’ll only accomplish greater things, if we understand how all the multiple facets of asset discovery fit together. And yet, discovery is only one piece of the overall picture—mostly from the past or what already exists in our environments. To provide more color and perspective, you need to add current and ongoing data that comes from day-to-day asset management activities. You also need to add more data with advanced patterns to visualize or forecast the future and future-proof your asset systems by helping you manage additional complexities.

Once you have a solid discovery practice in place and can see what you have, then model what you need to do with your assets, and—let’s workflow it!