Is Automated Infrastructure Documentation Practical?

David Cuthbertson, Square Mile Systems

With the increasing complexity of modern ICT infrastructures, the use of automated discovery toolsets is often seen as the only way to regain control. In practice however, many organisations have found that auto-discovery packages and management systems are oversold and expectations set too high. This white paper covers the practical use of toolsets and suggests approaches to help ensure that the ICT infrastructure knowledge base takes advantage of automation where practical.

Some typical questions we are often asked, with some simple answers.

Q. Is it possible to have an automated documentation system that doesn't involve technical teams having to input data manually? Answer : No…

Q. If I follow best practice guidelines such as ITIL for configuration management, does it mean I will have the right information to manage change in my infrastructure? Answer : No…

Q. Will an automated desktop asset management system make me compliant with licensing requirements. Answer : No…

It's easy to be negative, though it would be more useful to identify why there are gaps in understanding, expectations and delivery capability. What real, tangible benefits can be delivered?

Why do we want automation of our infrastructure documentation?

Perhaps the best way to answer this is to look at reasons why we don't manually document the infrastructure, although there are many benefits in sharing knowledge. Why is it so difficult to get operations and project teams updating a knowledge base when they install or change servers, routers, software, etc? Do we have to buy expensive toolsets because we can't get people to follow some simple rules or procedures? Maybe improved management practices could be the answer, so no toolsets are needed. Any form of automation will always require some processes to be developed and policies to be communicated. Better management practices are inherently part of the success criteria for any infrastructure documentation.

Each technology area (servers, networks, desktops, software, etc.) normally requires it's own set of specific documentation for project and operational needs. In addition there are teams, or functions which span technology and geographic areas (service desk, change management, security, finance, and business continuity) where the "big picture" is often required. Practically, this wider view cannot be maintained separately from the technology areas. If we summarise business needs and the potential value of infrastructure documentation, maybe the reasons for documenting become clearer and the role that automation tools can play.

If we had good, accurate infrastructure documentation we could probably:

- Reduce the cost of site surveys by project and change teams
- Enable better impact analysis of technical and business changes
- Be confident our infrastructure is secure
- Identify causes and recovery capability much faster when faults occur
- Negotiate better with suppliers on warranty, maintenance, support and licensing
- Reduce the cost of maintaining continuity and risk management plans
- Have more flexible support arrangements with local staff, central teams and suppliers
- More easily implement frameworks such as ITIL, BS15000, BS7799, PAS56
- Reduce SLA response times for provisioning of services, or responding to change requests
- Ensure accurate billing both internally and externally for services

Even if you agree with only some of these, why do we not maintain a structured knowledge base today - manually or automatically? Our experience is that both technology and business practices have constantly been evolving, with the value of operational best practices only recently being recognised in ICT. To reap the benefit of automated knowledge, we need to define what information is of real value to give focus. Or, in other words, work out your process needs first, then assess whether toolsets are delivering against them. This is often the first stumbling block, as it requires definition of strategic change with measurable objectives and ownership, which normally only senior managers have the skills and authority to define. If senior management involvement is limited then the risk of failure of any initiative increases, as their sponsorship is needed when developing internal processes.

What information do we really need?

Depending on the change desired, different types of information are required from the knowledge base. Planning change requires knowledge of existing capacity, dependencies , schedules and performance. Alternatively, managing an incident better often requires ownership to be identified, along with technical data about components and logical relationships. As a general rule, automated toolsets often provide technical data (server disk space, memory, etc) which can be overwhelming and sometimes inconsistent. What the toolsets cannot provide is the business information; the server doesn't know when it was bought, where it is, when the warranty runs out, who uses it, or what business function it is supporting. There is an instant conflict of interests between technical groups who would use the knowledge base – to some it is not detailed enough, to others it is too specific and doesn't show the bigger picture.

Some examples may help to illustrate the point. A desktop auditing tool was purchased and deployed to help a company speed up its adds/moves/changes process and also to help reconcile licensing issues. The reasoning was that, with detailed knowledge of the desktop configuration both needs would be satisfied. The end result was that neither benefit was met for the following reasons.
1. The adds/moves/changes planning required knowledge of cabling connectivity and other devices co-located such as phones which could only be gained from a site visit.
2. The audit tool would find lots of executable files on a desktop, relying on a translation table within the auditing software to identify versions of Word, Excel, etc. The translation table didn't cover bespoke and mainframe software, plus it couldn't cope with package units (ie MS Office).

At Square Mile we tend to split information further into that which is easily managed using databases, and that which is best managed using diagrams such as logical, system or process relationships. You start with the database and then complement it with visual representations of cabinet layouts, network diagrams, domain structures, application data flows and so on. As shown in Figure 1, the visual representations cannot be easily automated because you create a picture to suit a need. An analogy with houses, streets, roads and towns is useful as planning a trip would use a road map for high level routing, but a street map at the destination to know where a house is. Knowing the end to end path where users are having response problems assumes you have diagrams which combine user, host and infrastructure components.

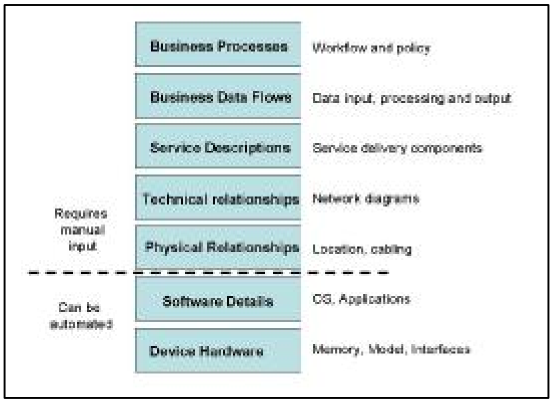

Figure 1 Presenting Device Information

What are the Known Problems with Automated Discovery Systems?

Every automated toolset has benefits, as well as drawbacks. So before committing to any system these are a number of factors to consider. Each question is asked due to prior experience – note that a **single** negative answer may totally preclude the use of any toolset.

- Can the toolset access all parts of the infrastructure consistently?
- How do technology specific toolsets (desktop, network etc.) use common references?
- What happens if a component is not detected – is it deleted, or flagged as undetected?
- How easy is it to put in equipment which is switched off, or manually inventoried and stored in spreadsheets?
- Do discovery agents have to be re-installed with operating system upgrades or service packs?

- Has the device information found automatically, been verified as being accurate across a range of hardware platforms?
- Where network discovery relies on unique identifiers, what happens when a device is connected on different media eg. LAN, wireless, remote access, or a LAN card is swapped
- Will the manufacturer or supplier verify that the discovery process will not adversely impact devices or services.
- How can the discovery process avoid auditing or detecting devices which are of no interest on the network? eg. Loan equipment, demonstration systems, non-company devices.
- Can device information be associated/grouped and reported in grouping? eg. Departmental, owned/leased within the system.
- If a device is replaced by a faster model using the same name/IP address, does the previous device details automatically get updated or erased?
- How is collected information linked to other systems so that automated updating is controlled and manually input information is verified?

The list could go on, plus you can add your own examples of servers found with zero memory, laptops with 100Mb of disk space, rather than 100Gb, routers with 50 network interfaces but only 10 external connectors. In practice, we have found that automated discovery systems can be of real benefit if you have no initial information, though you end up with a new problem of the discovery database increasing as deleting devices is normally a manual task. In one case we found that an organisation had 500 desktops, but 860 in the audit database. It was so far out that no-one trusted the information and verified everything manually, negating the whole point of buying the tool in the first place. Figure 2 illustrates this. The lesson to learn is that you will still have to undertake manual verification checks, even with an automated system.

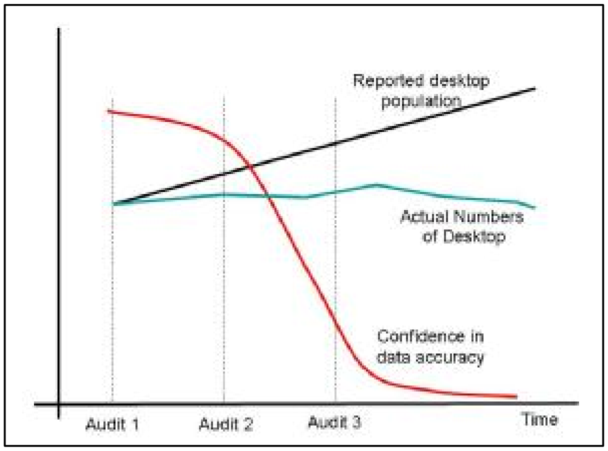

Figure 2 Audit Database without verification process

How should we Integrate automated discovery systems?

As with many simple questions, the answer may be complex. If you have no information, then the introduction of an automated discovery system gives immediate benefit. It's accuracy may be questionable, but it is still an improvement. If you have only an audit type approach of checking once a year it can help. You will typically only discover data about the devices, along with other technical information like network routing tables. Business related information and dependencies will have to be done manually.

If you already have controls already in place, such as a CMDB, help desk inventory or other data repository, then the discovery data may update the technical information to the latest state. There are the control issues of changes in devices, unique references and authority to delete, which will typically be handled by an individual tasked with maintaining the system(s). The same person can handle the maintenance of the business related device information. In practice, this needs to be thought out as otherwise you end up being overwhelmed with data as each discovery system often overlaps and duplicates. As an example, a network management system discovers all IP devices. The server management system discovers all Wintel servers and gives greater depth on server details. If you added a new server to the network, you have two sets of information to update a central repository – do you add both, or combine with one being a master? When you add a Linux server, it becomes more interesting. In smaller environments these issues can be manually worked around, in larger installations you can go round in circles.

For a coordinated approach, we recommend that discovery tools are treated as a source of information, but the maintained data repository is kept separate. This database has processes to support it, along with policies and responsibilities. Someone must own the information and its accuracy, with discovery systems being used to support/verify technical device details. This way we can have a central system that copes with normal asset information such as status, (live, spare, repair, faulty, etc.) and also enables the organisation to work around the inconsistencies of discovery tools. Some organisations use the repository approach to push updates out to monitoring and discovery tools when a new device is added. When a new router is purchased, it is entered into the repository, which then adds its details to the discovery and monitoring systems to ensure that the device build details are captured and performance is monitored. The discovery tools are used to verify and capture detailed technical information, but not relied upon as the definitive source of knowledge ie. the manual system is supported by automated discovery – rather the automation requiring manual support.

Another aspect of automation concerns the linking of databases together, to save re-entering of information which has been collected by either manual or discovery methods. This approach offers the most in terms of business benefits, but also requires detailed definition of working processes and ownership. In the service provider community this might be known as an OSS (Operational Support System) and requires a proper application development programme. With any integrated system composed of multiple data repositories, the risk of failure increases without the appropriate resources and sponsorship.

In conclusion

This white paper has looked at automated discovery tools from an overview of controlling assets, their configuration information and the processes that need detailed knowledge of IT environments to be effective. In smaller infrastructures, or single technology departments, automation can deliver benefits, though only providing a subset of the information you often need. Manual reconciliation is enough in many cases to tie up the loose ends. In larger environments, where you typically need to get end to end understanding within your asset and systems documentation, it is imperative to ensure that manual processes and information sets are defined before expecting any automated toolsets to deliver significant long term benefits.

The Author

David Cuthbertson is a founding director of Square Mile Systems, a UK computer services company based in Cirencester, England. He is an industry speaker on best practices and applying configuration management techniques to ICT infrastructure. He is also chairperson of the BCS Service Management Specialist Group (SMSG), as well as chairing the Board of Governors for the Academy of IT, a further education initiative developing vocational training for new entrants into the IT industry.

Worldwide Headquarters (Europe & US)
Square Mile Systems
3 Church Street, Cirencester
Gloucestershire, United Kingdom GL7 1LE
Tel +44 (0)870 950 4651 Fax +44 (0)870 751 9268
www.squaremilesystems.com

Asia Pacific (including Australia)
Tel +61 (0)417 231726