How Quantity and Diversity of Data Sources Determine What's True... For Your CMDB and Beyond

EMA™

*IT & DATA MANAGEMENT RESEARCH,
INDUSTRY ANALYSIS & CONSULTING*

Table of Contents

EMA

Executive Introduction

While it's tempting to cling to a single "trusted source" for discovering IT assets and their service-related interdependencies, too often it creates a false sense of security that can lead to disappointments, failures and sometimes even catastrophic results. This is because optimizing IT assets and managing IT services—whether for change and configuration, or for performance and availability—typically depends on many dimensions beyond the capabilities of a single discovery or inventory tool. Moreover, IT hardware and software infrastructures are varied and complex, and no single source is optimized for the full range of technology types, brands, and versions that exist across large, enterprise distributed environments.

This white paper looks at research from EMA, as well as from Blazent, to highlight the requirements for reconciling and normalizing multiple sources in support of superior IT service management effectiveness. It will also include comments from real-world deployments to underscore both requirements and best practices in building a system to support a CMDB or CMS capable of weathering the dynamic pressures of contemporary IT and business environments.

> This white paper looks at research from EMA, as well as from Blazent, to highlight the requirements for reconciling and normalizing multiple sources in support of superior IT service management.

A Single Source of Truth?

Since EMA began tracking CMDB deployments in 2004–2005, it became clear that interest in a Configuration Management Database (CMDB) back then was driven by multiple factors, which EMA called the *CMDB's two parents.* This dual parentage continues to this day, and remains a source of much confusion in the industry.

The *first parent* is of course the IT Infrastructure Library (ITIL), which first coined the term CMDB and defined CMDB requirements very strictly in terms of roles and process objectives within IT. It remained cautiously behind the scenes, by contrast, when it came to architecture and discovery requirements. Now central to understanding the CMDB, is ITIL v3's definition of Configuration Management System (CMS), which may include one or multiple CMDBs, in *service transition.* Whether it's a single CMDB, or a part of a larger CMS, it should include information about incidents, problems, known errors, changes and releases; and may contain data about employees, suppliers, locations, business units, customers and users.
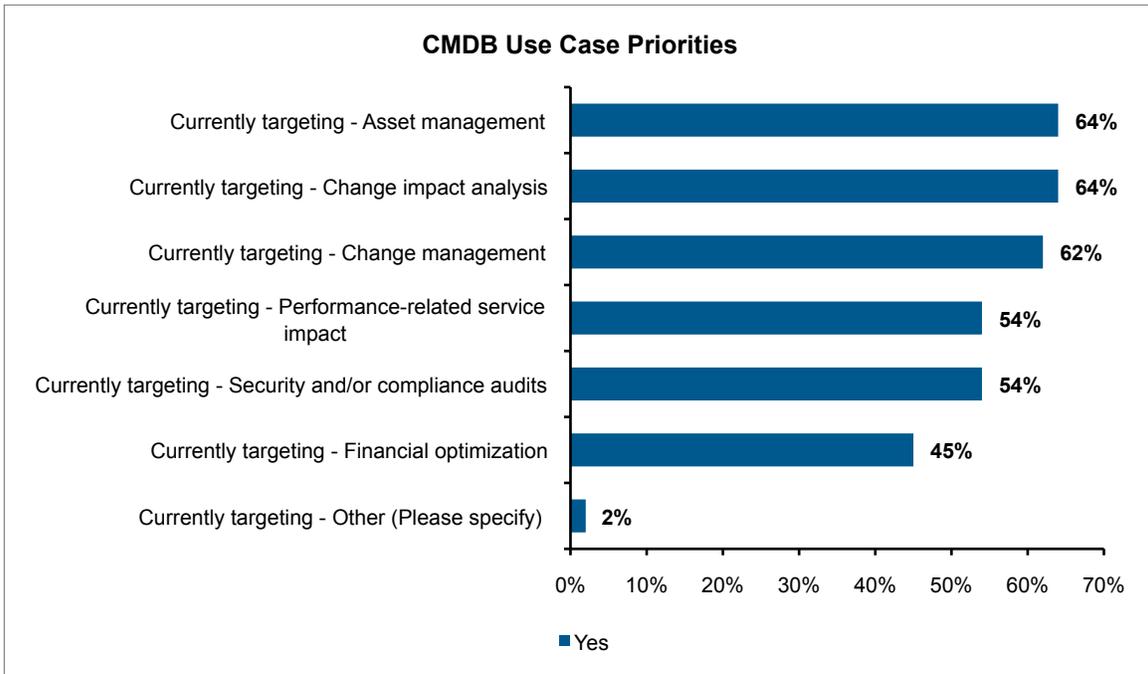
The CMDB's *second parent* is primarily architectural. It is a driver associated with the need to assimilate different management investments in a manner that supports superior levels of analytics, decision-making and automation. It depends on increasingly dynamic and effective ways for importing and reconciling data from a wide range of sources, from spreadsheets to other databases to other discovery and management tools—deconstructing the siloed manner of management tool design in the past to support cross-domain requirements to analyze and optimize information.

Leveraging Multiple Data Sources for Completeness, Currency and Accuracy

One of the ways to better understand the need for breadth of data coverage in an effective CMDB/CMS, or other strategic deployments, is to understand that any given source is, invariably, linked to one or more use cases. Discovering "what's out there" is valuable at once to asset management, change

EMA

management, and performance management—each with their own set of tools targeted for a per use case basis, but often producing huge amounts of confusion, inefficiencies and redundancies if not optimized as a whole.

CMDB Use Case Priorities

Use Case	Percentage
Currently targeting - Asset management	64%
Currently targeting - Change impact analysis	64%
Currently targeting - Change management	62%
Currently targeting - Performance-related service impact	54%
Currently targeting - Security and/or compliance audits	54%
Currently targeting - Financial optimization	45%
Currently targeting - Other (Please specify)	2%

■Yes

Figure 1: Multiple use cases require a wide variety of well-reconciled insights. The average CMDB/CMS deployment above shows 3.4 use cases

Moreover, capturing interdependencies, the very heart and soul of good CMDB design, depends on being able to look at and understand a given CI from many different points of view through many different toolsets, spreadsheets, databases and other sources. The sad truth is that no single discovery tool, no matter how good, is perfect for all occasions across all forms of infrastructure, software and applications.

But the challenge of reconciling these insights goes beyond sheer complexity and volume (EMA documented one mid-tier enterprise in the manufacturing sector with multiple geographically-dispersed data centers, with more than 500 brands of monitoring tools!) it also gets political. It's fine to say that you've assigned tool XYZ as a trusted source and set up policies to that effect, but how do you really know?

> The sad truth is that no single discovery tool, no matter how good, is perfect for all occasions across all forms of infrastructure, software and applications.

Most current solutions are optimized for setting policies and some minimal analysis on CI redundancies, but can't come close to resolving the frequent and all but inevitable debates between stakeholders shouting, "My tool is better than yours!" This is also shorthand for saying, "I'm more accurate and better than you are, and therefore more important to the organization." In other words, trusted source disputes can quickly degenerate into veiled debates about job security. This can, needless to say, inhibit the very commitment to utilize multiple data sources for better insight and breadth.

This can also impact currency, as insights into how infrastructure, software, services and monitoring capabilities change can become obscured by source fragmentation. EMA is increasingly seeing priorities to update CMDBs in near real-time given trends like virtualization, cloud, and consumer-driven needs for service performance optimization.

EMA

Breadth of Data Coverage

When evaluating CMDB/CMS capabilities for strategic, cross-domain initiatives, EMA asks the following:

Range of Discovery: Can you natively, or through third-party integrations, support discovery for network (layer 2 and/or 3), systems, applications, application components, third-party applications, Web and Web 2.0, storage, database, desktops, mobile devices, and virtualized environments? Can you discover configurations for the above?

What Types of Sources Can You Discover? Ideally, this should include Management Data Repositories (MDRs), text records, Excel or other sources imported from your own portfolio and third parties such as service catalogs, service desks, performance management tools, security tools, asset management tools, and other configuration management tools, etc.

But what does this really mean when you really lay it out? Just in terms of discovery, EMA looks for the following:

Asset inventory integrations
Network discovery (Layers 2/3)
Systems discovery
VM and virtualized system discovery
Broader support for mixed (network and or storage) virtualized environments
Mainframe
Storage
Peripherals and desktops
Applications "out of the box"
Applications – custom
Virtualized applications
VDI environments
Application-to-infrastructure interdependencies
Infrastructure-to-infrastructure interdependencies
Application-to-application interdependencies
Security-vulnerability-related discovery

Figure 2: Range of discovery—domains—as per the EMA Radar™ for Application Discovery and Dependency Mapping

Moreover, these are in fact general domain targets—a deeper dive, just in virtualization, might reveal the following as examples:

Clustering
VMware ESX
OracleVM or Sun xVM
Citrix Xen Server
Linux Xen (SUSE Red Hat, etc.)
Microsoft Hyper-V
IBM PowerVM
IBM Z/VM
Linux KVM
Parallels Virtuozzo

Figure 3: A deeper dive—using just virtualized environments as an example

Finally, there's functional breadth so critical for capturing meaningful interdependencies. The sample below is targeted, albeit at a still high level, at asset management.

Basic inventory
Software asset management (custom)
Software asset management (COTS)
Software license management
Hardware usage and maintenance
Other asset-related compliance reporting
Green IT initiatives
Analytics for capacity planning and optimization
Lifecycle asset management and planning
Service portfolio planning
Virtualized systems tracking

Figure 4: Effective asset management functionality – breadth requirements for effective CMDB and other strategic integrations require insights across process and function as well as unique domain hardware and software insights.

Accuracy

Another way to understand breadth is in terms of accuracy. Multiple data sources will also provide complementary insights into how the pieces of a CI jigsaw puzzle fit together. The example below is what EMA looks for just in agentless discovery.

IP ping sweep
Layer 2 discovery
WMI
MAC address
SNMP statistics
Port-level interrogation
Flow-based analytics
Observed transaction analysis
Code analysis

Figure 5: Agentless discovery from multiple tools will provide complementary insights into what's true and what's not with a CI, and with CI inter-relationships

Another set of parameters might include deeper insights into specific devices—in this case, servers. Below is a partial list from EMA evaluations not including detailed configuration data:

Systems "users" of database transactions
Systems CPU
Systems manufacturer
Systems manufacturer name
Systems product name/version
Systems memory
Systems network interfaces
Systems files (name, description, size and path)
File modification time
File extension
Directory file time
Systems running processes
Executable type
Systems daemons

Figure 6: Different sources capture different details while looking at the same device or CI. They also offer suffer from different naming conventions, making reconciliation all the more challenging

Perspectives from EMA Consulting and Research

EMA has done extensive research and consulting in both CMDB adoption, and in support of strategic, cross-domain initiatives outside of CMDB adoption. As an example, one of EMA's clients, a large company in the financial services sector based in North America, documented no fewer than **72 different data sources** in planning their CMDB!

The comments below come from both research and consulting reflecting various CMDB and other initiatives.

Struggling with Redundant Data

"We estimate that currently we are sitting on about $20 million in redundant, unreconciled data from our many multiple data sources. SMS is used in some places. Network does its own discovery. Mortgage has its own. Credit uses its own. So we have lots of redundant software and processes at work.

"We do 120 asset extracts a month. But we have no consistence in optimizing our data. We also need a solid inventory of applications. But there are no common definitions, no common vocabulary, no place to meaningfully reconcile and store the data. Management investments are domain-focused. There appears to be little consideration for integration when new tools are introduced into the environment; while, most asset data is stored in an 8000-line spreadsheet. As a result, there's a disconnect between asset and auto discovery. Auto discovery is good, but manual feeds from asset show a gap that needs reconciliation."

> *"We estimate that currently we are sitting on about $20 million in redundant, unreconciled data from our many multiple data sources."*

Breadth of Stakeholders

"We have 50 direct users and about 9,000 indirect users, as our CMDB supports the entire business.

"I would say that two-thirds of our 300-person IT organization are currently stakeholders of the larger CMDB/CMS system either directly or indirectly. And interest continues to grow as more and more people in IT begin to realize and see value in accessing information that they didn't know was there before.

"We have about 300 different types of stakeholders defined including business executives, applications management, desktops, servers, mainframes, merger recovery services, facilities planning, and ITIL process owners—just to name a few."

The Politics of Unity

"The challenge is still making the organization use CMDB. Today we still have some teams that use EXCEL CMDBs and don't want to use a discovery tool because they don't trust it.

"We weren't going to win the political battle to get people to give up their tools for just one centralized solution. So we needed a CMDB that facilitated and reconciled data from many different sources, many different brands."

Perspectives from Blazent Deployments

Blazent recently shared histories with EMA from seven recent enterprise deployments that averaged **18 different data sources per customer**. Moreover, when EMA did a tally, unique data sources exceeded 100 across all seven environments. This once again underscores requirements for achieving accuracy, breadth of scope, and breadth of functional values through collecting data from multiple, often imperfect and domain-centric, separate sources.

Below is a telling quote from a Blazent deployment interviewed by EMA earlier this year.

"We have 56 different data feeds – everything from systems inventory, desktop inventory, anti-virus, Active Directory, encryption, software capacity reports, software distribution, lease schedules, etc. Moreover, Blazent allows us to categorize and manage our CIs logically. For instance, devices connected to the network that can be routinely discovered can be analyzed separately from end points which may or may not be connected to the network at any given time, and training or development devices that are only connected intermittently. Blazent has vastly improved our visibility and efficiencies."

> "We have 56 different data feeds – everything from systems inventory, desktop inventory, anti-virus, Active Directory, encryption, software capacity reports, software distribution, lease schedules, etc."

Conclusion

An effective game plan for establishing either a CMDB/CMS, or another strategic initiative aimed beyond a single, narrow domain should have the following characteristics.

- Breadth in terms of interrelated domain perspectives.
- Breadth in terms of types of data sources.
- Breadth in terms of interrelated functional perspectives (e.g. inventory, license insights, configuration, utilization, etc.).
- Breadth in terms of data discovery and collection technologies and KPIs.
- Well thought-out support for multiple stakeholders and roles.
- A system for automating the process of reconciling and normalizing data from many multiple sources.

Finally, it's important to realize that just because a "trusted source" is trusted by someone, that doesn't necessarily mean it's complete or accurate. Beyond mere policy setting, "trusted source analytics" is critical to optimizing what might otherwise become an unruly chaos of options.

The benefits for making these investments, on the other hand, can be tremendous. EMA's report on ROI and metrics associated with CMDB/CMS deployments clarifies this in more detail, but the bottom line is that optimizing discovery, monitoring, spreadsheet and other data sources provides *the* best foundation for enabling cross-domain efficiencies, and allowing IT to evolve from a reactive, cost-averse, to a proactive, value-driven business contributor.

About Blazent

Blazent is the world's most widely-used IT Data Integrity Engine. Built on patented algorithms developed over the last decade, Blazent's cloud-based engine aggregates, reconciles and consolidates IT data to guarantee continuous accuracy, and to optimize IT management & operations. Global 5000 executives rely on Blazent to ensure effective governance & compliance, mitigate risk, control costs and support major business transformation. As the gold standard for IT Data Integrity, Blazent empowers the business of IT.

Blazent's CMDB Accuracy Solutions for BMC and ServiceNow deliver an automated, user-friendly data reconciliation process that enhances CMDB data accuracy. The solutions leverage innovative two-way connectors that allows users to identify CMDB data quality issues within Blazent and then initiate a service request to fix them with just a click of a button. These requests are sent electronically to the BMC Remedy or ServiceNow platforms, routed through the appropriate Incident, Change or Problem processes, and successfully resolved. Regardless if you are trying to improve an existing CMDB, switching CMDB vendors or implementing a CMDB for the very first time, Blazent can help ensure your CMDB is populated accurately and stays accurate over time. Blazent's CMDB Accuracy Solution was approved as a Certified Integration within ServiceNow's Technology Partner Program in August 2013, and is currently undergoing certification within BMC's Technology Alliance Partner Program.

For more information, visit www.blazent.com or follow Blazent on Twitter @Blazent.