

The Njini IAM Suite

February, 2006



Until very recently, the automatic classification and control of information was considered an exotic IT luxury. By 2008, we believe that Information Classification and Management (ICM) will be a non-negotiable necessity for enterprise IT. The rise of ICM is directly correlated to the positive feedback loop created by two trends: data growth and storage infrastructure complexity. Both Data growth and technical complexity have exploded and fed off of each other in recent years, leading to a range of new innovations. As a result, we now have a vast proliferation of storage tiering, archival and ILM initiatives that address all manner of customer challenges. However, the little noted fact is that any next-generation response to today's data tsunami absolutely requires some *intelligence* at its head in order to ensure a successful deployment. Creating new complex storage architectures is never going to eliminate data complexity. For this reason, we have been very bullish on the ICM category, as it provides the tools, approaches and automated intelligence needed for fine-grained controls atop a storage infrastructure. Within this category, Njini Inc. stands out for its unique architectural approach and breadth of functionality. In this profile, we will provide an overview of the Njini IAM Suite and discuss the impact that this software platform can have on how enterprises finally gain control over their data growth and storage complexity challenges.

Information Overload

It has been well discussed how data growth rates are raging out-of-control. Few readers would doubt this state of affairs, and we can verify that we routinely see 40% to 70% growth rates now depending on the type and size of customer environment.

Ironically, the range of solutions available to address this "data tsunami" is itself a force bordering on the out-of-control: ILM, HSM, NFM, CAS, D2D, storage tiering, migration tools, online archival, to name but some of the most common remedies prescribed to address storage growth today. All of these technologies are powerful and have their place in the next-generation data center, but they beg a key insight that has been often

over-looked: Adding new complexities to the storage infrastructure does nothing to resolve the *informational complexity* behind storage growth.

Without some manner of software intelligence that connects the business's informational value to the data that is being stored, no manner of new storage innovations at the infrastructure level stand a chance of alleviating the enterprise IT team's data challenges. What is needed is "brains" at the head of the new storage infrastructure. Fortunately, there is a new product category called Information Classification and Management (ICM) that has emerged to address this need, and we believe the technologies in this category can go a long

P R O D U C T P R O F I L E

way to solving storage complexity, and making new storage investments pay-off.

Why Classify Anything?

We have had many end users ask us why they should deploy specific classification and control solutions for their enterprise content. In short, why classify anything at all? In response, we always point to four advantages that derive from the ICM category and enterprise information classification. We have itemized these advantages below.

ICM Advantage: Gaining Control

The optimal time to get control over data and valuable information about that data (i.e. metadata) is when it still maintains an immediate application or usage context. The answer is not to wait until *after* it has been divorced from that usage context in a data repository. The best ICM tools are designed to enable rapid, immediate decisions based precisely on this principle of immediate classification and control. Some offerings actually provide these controls in real-time, as we shall see. By deploying an ICM solution that enables immediate classification of content, enterprises derive a range of operational efficiencies that are otherwise impossible.

ICM Advantage: Stop Duplicate Data

The enterprise is riddled with redundant data sets that reduce utilization, performance, and overall ROI. This is especially true in distributed and high growth environments. A good classification tool can immediately identify and eliminate redundant data through single-instancing, while maintaining the original informational value of the object.

This has a direct impact on storage economics.

ICM Advantage: Smart Data Placement

A major benefit from ICM tools is the intelligent and automated placement of data within storage repositories. Rather than migrate data later, a tiering deployment can be significantly optimized by selecting the appropriate targets based on pre-established policies. This means a significantly more efficient storage deployment, with less need for manual human intervention and error.

ICM Advantage: Link Business to Tech

A good ICM tool will enable the enterprise to link a range of business level initiatives directly to the classification process, and thereby gain economic and operational benefits from this connection. The most obvious areas where we see this linkage taking place today include regulatory compliance and information security initiatives, in addition to ILM and tiering projects.

As these advantages demonstrate, ICM is a technology with a potentially far-ranging impact. Classification *does* matter, and is in fact the gateway to an entirely new level of enterprise data control. To understand just how broad this impact can be, we now turn to explore the unique approach to ICM taken by Njini Inc. and that company's IAM Suite software platform.

Meet Njini

One of the emerging players of note in the ICM space is Njini Inc. The company provides all software-based solutions for

PRODUCT PROFILE

classification and control of unstructured (file) content. Their flagship product, IAM Suite (IAM stands for Information Asset Management) is comprised of several interconnected software elements that together provide an entire range of enterprise ICM functionality.

End users should conceptualize Njini's IAM Suite as a solution for providing real-time classification and management across the entire file infrastructure of their organization. First, it will be helpful to understand the architecture behind Njini IAM Suite, so we can better understand its implications for an enterprise data center.

IAM Suite Components

Njini has built their solution to be highly modular in deployment. All of the software modules for IAM Suite can be deployed on a wide range of industry standard server configurations, running Red Hat ES 4.0. The entire Njini IAM Suite is always deployed between the end clients and the networked storage infrastructure. This is an in-band solution, eliminating the need for any client-side agents or device drivers. Accordingly, access to the Njini solution takes place through a standard CIFS or NFS interface. Designed to be totally transparent to the end client machines, it leverages the same authentication mechanisms of their CIFS shares or NFS exports. The offering can operate equally well with direct-attach or FC SAN environments.

The Main Muscle: njiniENGINE

Njini refers to each core classification node in a deployment as a njiniENGINE. To achieve

very high scalability (2,000 users per "engine"), the njiniENGINE can be clustered across distributed or unified server environments, leveraging whatever node to node transport protocol the user has deployed (e.g. TCP/IP, RDMA, or shared memory). The Njini database architecture leverages a POSIX file system to surface transaction-coherent views across all servers, regardless of the storage device on which they reside.

So, what is actually stored in these Njini databases? All of the user metadata regarding the infrastructure's file contents are maintained within, as well as policies regarding how data should be stored and accessed. This includes attributes such as ownership usage and access history, timestamps and security settings, and instructions for what storage resources given data sets should reside on. Njini can reside on either Postgres or Oracle 10g database deployments.

Once deployed, the NjiniENGINE controls the actual classification of user data, all conducted in real-time. Using a process they call "Identity@Origin", Njini integrates its file controls directly with existing Active Directory to ensure immediate file controls that are seamless with the operating environment.

It is important to understand that all of Njini's classification and control capabilities derive from the actions of their njiniENGINE nodes. Once a user has been authenticated within the Njini system, all of their data can then be controlled and managed across the infrastructure. Depending on what the user

P R O D U C T P R O F I L E

wants to achieve with their IAM Suite deployment, the njiniENGINE can enable several helper modules that the company provides to achieve specific business or operational goals. To understand the broad impact that Njini can have, we will now turn our attention to these modules.

Data Duplication: njiniENCOUNT

One of the most valuable elements in the IAM Suite is the module called njiniENCOUNT. This is an application that leverages the njiniENGINE to remove unwanted duplicate data. In our opinion, this is the module that most end users will use as the centerpiece of their ROI justification for purchase. Recall that because Njini operates in real-time, with the njiniENCOUNT module, the IT manager can essentially eliminate duplicate data *before* it even becomes static data taking up capacity in filer servers and storage resources. In a world where file content can often multiply by 4-8x without anyone explicitly noticing, this is a major efficiency gain.

Data Placement: njiniENROLL

Another major value that can be derived from the njiniENGINE sitting in the data path is the ability to conduct real-time bi-directional placement of data based on policies. With this njiniENROLL module, IT teams can define policies based on a wide range of usage attributes, including content aware notifications, access patterns, modification histories, or storage architecture changes. Dual processing then occurs, allowing either explicit initial file placement anywhere within the storage architecture at the time of creation or later when a given attribute changes. In either

case, the njiniENROLL module can then facilitate file level migration between physical data stores. This enables perpetual updating of storage tiers based on the latest available operations. We believe this could have a significant impact on how enterprises think about ILM deployments, as it points toward an era of beyond bulk migration, where true intelligent automation is actually possible.

Information Security: njiniENFORCE

The next module available in the IAM Suite is a security and access control capability called njiniENFORCE. Recall that the Njini solution enables the IT manager to establish and automatically authenticate user access rights based on any number of customizable variables. When combined with Njini's content-aware classification capabilities, it is then possible to enforce which users have access to what content, based on highly specific criteria (e.g. keywords, document types, content associations, or any user-defined parameter.)

We see the immediate application for njiniENFORCE to be both regulatory compliance and informational security. In both of these usage cases, fine-grained control over individuals and group access rights is absolutely critical. Historically, this level of real-time, customizable user control was impossible, especially in a heterogeneous or distributed deployment scenario.

Content Search: njiniENQUIRE

The final module within the IAM Suite focuses on retrieval and access of the content that has been indexed. With njiniENQUIRE, users or applications can conduct content

P R O D U C T P R O F I L E

and metadata search queries against variables, attributes, policies or keywords. Based on these queries, njiniENQUIRE then automatically retrieves all relevant files, directories or shares that match the query. This manner of “search and control” data management has significant implications for data management. With the ability to access and control any kind of distributed unstructured content in this highly centralized fashion, IT teams can very quickly address formerly time consuming issues such as individual user-data restores, migration planning, information audits, and legal discovery requests.

Where Njini Adds Value

We have now taken a quick tour of the Njini IAM Suite’s architecture and its key usage cases. It should be clear that the highly modular architecture of the Njini solution lends itself to solving a wide range of information management challenges. However, at the end of the day, how does one go about justifying an investment in advanced classification technology? What kinds of value should be measured?

To answer that question we have outlined several impact areas that we encourage IT teams to consider as they evaluate whether or not Njini IAM Suite makes sense for them. Based on our conversations with end users, there are 3 clear-cut areas where Njini IAM Suite is going to add measurable value.

Impact Area: Storage Efficiencies

There are several ways that Njini can create measurable, quantitative improvements to the storage environment. First and foremost,

because the IAM Suite can remove duplicate data at the source, before it becomes redundant data, it improves storage capacity efficiencies. This means less over-provisioning and a higher utilization rates for existing resources. We believe that this can have an average increase of 25-60% for a typical enterprise customer’s file content, depending on workflow and workloads.

Another potential efficiency results from Njini’s ability to move content to the appropriate tier of storage at the time of *initial* data placement. This means that the correct tier of storage is using the right data at all times. Teams can even avoid time-consuming bulk migrations if they craft appropriate methodologies around Njini’s classification policies. These “ILM” capabilities resulting from smart placement can prevent static content from taking up storage space, thereby increasing utilization rates and performance gains for the entire infrastructure. Because of this functionality, we estimate that Njini customers will be able to register between 25% and 40% gains in ROI for unstructured data storage resources.

Impact Area: Management Efficiency

A range of human gains can be had from automated user security and access control management, and end-user data retrieval requests. Most importantly, we see a range of *significant* storage management gains to be had with Njini (e.g. less touches for migrations, provisioning, capacity balancing, and performance tuning.) Based on the specifics of the customer’s environment and applicability of Njini usage cases, these efficiencies should be factored into a

P R O D U C T P R O F I L E

customer's total ROI argument along with hard technology costs savings.

Impact Area: Mitigated Risk

Many of the companies we speak with are very interested in ICM technologies because they need to mitigate their risks of unwanted information exposure. Often, these concerns are driven by regulatory compliance initiatives, but they need not be. Across industry verticals, we increasingly see security officers extending best practices for data controls to cover their unstructured content. This is driven by an understanding that unstructured content stores are host to a range of hard-to-detect security infractions. To this end, because Njini automates real-time, infrastructure-wide usage control policies, even across multiple geographies, it provides enterprises with a powerful tool for centralizing security and usage policies. This is an impact area very unique to the ICM category today: Without knowledge of the relationship between content and usage, comprehensive user, access and security policies are not possible for unstructured content.

Taneja Group Opinion

The advantages that result from the proactive classification of enterprise information are becoming increasingly well understood. For any company that is seriously pursuing comprehensive ILM, a storage tiering plan, or implementing access and usage controls for compliance and security, ICM should constitute the critical "brains" of their deployment.

We are often asked how "mature" the ICM space is in 2006. To be sure, it is still an emerging technology category, with a wide range of innovation and integration still ahead of it. However, it is definitely ready for "prime time", and we hear this opinion echoed across the end user community. In fact, based on the real ROI and management gains that customers are achieving with today's ICM, we have no doubt that these technologies are on their way to an established role in medium and large enterprises over the next 18 to 36 months. During this time frame, any serious attempts to drive enduring storage infrastructure ROI improvements without ICM functionality are risking sub-optimization of their returns. It is simply too hard to see and control the vast amounts of unstructured content in the typical enterprise without some manner of proactive, intelligent and actionable classification.

For these reasons, we are very keen on what Njini Inc. has brought to market with their IAM Suite. This is the only dedicated in-band ICM technology on the market today, and that distinction provides the offering with a range of potential advantages including real time de-duplication, real time data placement, real time security and access controls. These are capabilities that their ICM competitors do not possess today. Looking forward, we believe that the Njini architecture opens up many new doors for classification technologies, including entirely new markets and productization opportunities.

From a partnering standpoint, we believe Njini should be attractive to a range of

P R O D U C T P R O F I L E

players in the networking and storage networking world that have previously been unaddressed by the ICM category (think of both network and storage switching vendors looking to augment their offerings.) Indeed, Njini seems to provide an important and natural bridge between network and storage intelligence, as a result of its in-band design center.

Regarding the reliability of the Njini IAM Suite's architecture, we believe that the company has thought hard about the mission critical specifications required to support ICM as an in-band approach (e.g. availability, distributed scaling, performance) without sacrificing modularity and extensibility of the offering. Most importantly, their customers seem to agree as Njini's early market traction confirms.

For enterprises that are serious about getting control of their information assets and pleasing their CFO, Njini IAM Suite merits serious evaluation. Njini Inc. is an innovative company with a unique approach to a complex problem. For these reasons, Njini makes our shortlist of companies to keep an eye on in 2006.

***NOTICE:** The information and product recommendations made by the TANEJA GROUP are based upon public information and sources and may also include personal opinions both of the TANEJA GROUP and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. All product names used and mentioned herein are the trademarks of their respective owners. The TANEJA GROUP, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors which may appear in this document.*