

EMA RADAR™ for Storage Intelligence

By Steve Brasen

ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) Radar Report

August 2017



The Scope of Storage Intelligence

The performance, reliability, and availability of IT services are entirely dependent on the infrastructure hosting them, and no infrastructure components are more critical to business productivity than storage units. Certainly, IT service availability would be impacted if critical servers or networks become non-functional, but these are recoverable by simply replacing the damaged components. A loss of a storage system, however, may result in the loss of critical business data, and this could have significant consequences for business performance and profitability. Further, workforces are increasingly dependent on remote access to business data in order to perform job tasks and achieve organizational goals. Poor storage performance can therefore directly limit workers' ability to support customer needs and respond to rapidly-changing market requirements. Recognizing the importance of a reliable storage environment, many organizations have invested heavily in storage components—grossly overprovisioning in order to provide some level of confidence they have sufficient IOPS and capacity to support business requirements. However, this approach does not guarantee long-term viability of storage investments, and is clearly not fiscally responsible for the business.

To ensure deployments are business-focused, the primary goal for storage management is to enable data ecosystems that are both reliable and cost-effective. Achieving this requires holistic visibility into storage use, performance, and status that is easily digestible by IT managers so informed decisions can be made on optimal implementations and configurations. Storage intelligence encompasses all of the management processes necessary for collecting, analyzing, and presenting information on storage resource performance and utilization. Many hardware vendors offer storage intelligence solutions that provide information only on their specific product line, while some software solution providers offer storage intelligence platforms that address a variety of physical and virtual storage resources. The output from effective storage intelligence solutions provides guidance to IT managers on capacity planning, service performance improvements, reliability assurance, and compliance achievement. EMA believes all storage intelligence platforms must provide functionality in three key areas:

- **Data Collection** – Status, usage, and performance information is collected from all storage-related devices, networks, and/or cloud-hosted services. Collection services can be either native to the solution or incorporate data collected by third-party management platforms. All records should be stored in a centralized data repository for easy access and analysis.
- **Analytics** – Comprehensive information across an entire storage ecosystem is far too complex and vast for administrators to regularly review and evaluate using purely manual processes. The employment of analytics allows data to be automatically correlated to strategically identify the specific elements or conditions that require administrator attention. Analytics can be used to identify historical trends, indicate out-of-compliance systems, determine opportunities for environment improvements, and create hypothetical models on the projected future state of a storage environment.
- **Reporting and Alarming** – Analytical results are presented in a format (such as a chart, graph, or status indicator) that enables rapid identification of status, potential problems, and service improvement opportunities. This information may be presented in dashboards or generated in ad-hoc reports. In the event an analytical process identifies a faulty or out-of-compliance storage element, alerts may be sent to administrators so they may receive prompt attention.

Storage intelligence capabilities are foundational to enabling comprehensive infrastructure performance improvements and are an essential component of a software-defined storage environment. It is simply impossible to continuously ensure high-availability, reliability, performance, and cost-effectiveness without holistic visibility across the entire storage ecosystem. In fact, the true power of a software-defined data center is revealed when application performance details are mapped directly to storage intelligence information to enable the dynamic placement of workloads in a way that is optimized to meet end-user and business requirements. Adoption of a storage intelligence platform, therefore, will not only help

provide immediate relief to exiting storage-related service problems, but will also set the stage for achieving long-term goals for enabling dynamic data center management and continuous service compliance.



Assessing the Storage Intelligence Market

To assist organizations in identifying storage intelligence solutions that will provide the greatest value to their unique business requirements, EMA evaluated the leading platforms available on the market today. EMA defines “value” as the ratio derived from the strength of a product set against its cost efficiency. Put simply, the more a company pays for a solution, the greater the advantages they should receive in terms of breadth of functionality and supportability. EMA’s review process began with the determination of critical storage intelligence features and capabilities. EMA used this list to establish evaluation KPIs that were ranked and weighted to correspond with the prioritized requirements by organizations that have adopted or plan to adopt a storage intelligence platform. The prioritization determinations are based on discussions with IT operations managers, survey-based research responses, and EMA’s own experience and knowledge of enterprise requirements and best practices.

From these KPIs, EMA established a minimum level of functional requirements to identify which management platforms qualify for recognition as leading storage intelligence solutions. Minimum requirements included providing support of all three principle storage intelligence functional elements (data collection, analytics, and reporting and alarming). EMA identified and reviewed dozens of vendors offering storage intelligence solutions. Ten of these vendors were selected as offering the most comprehensive storage intelligence support as defined by the established KPIs, and each was invited to participate in the in-depth evaluation process.

A detailed questionnaire on the capabilities, cost, and supportability of their respective product sets was submitted to each of the selected storage intelligence solution providers. More than 200 points of comparison were considered and all responses were carefully vetted for accuracy. EMA also conducted interviews with vendor customers to confirm product capabilities and indicate customer satisfaction with the product sets. Scoring of the vendor solutions was mathematically calculated by correlating available features, architectures, pricing, and capabilities with the predetermined KPIs. Some individual feature scores were adjusted based on firsthand customer experiences with the product sets. Final scoring of each product set was used in the product comparison charts and in the determination of award winners.

Characteristics of a Preferred Solution

The EMA RADAR™ report evaluation process standardizes the review of product sets in specific management disciplines by comparing vendor and product characteristics in five distinct categories: architecture and integration, functionality, deployment and administration, cost advantage, and vendor strength. Identified below are the elements EMA believes are indicative of an ideal storage intelligence solution in each of the primary evaluation categories.

Architecture and Integration

The ideal storage intelligence solution is a fully integrated platform with a centralized console interface, data repository, and reporting engine. Solutions may be hosted on physical servers, appliances, virtual servers, or public cloud environments, and should be easily scalable with expandable components, such as by leveraging database rollup capabilities. Data collection processes should employ a variety of different protocols (such as CLI, SNMP, WMI, MIB, SIM-S, OID, and SSH) to ensure the broadest possible opportunities for information acquisition. Additionally, data collection processes should have little or no impact on network and storage system performance such as by employing light-weight agents, agentless collectors, or passive network sniffers.

Broad data collection capabilities may provide comprehensive details on the storage ecosystem, but the information is far too numerous for human administrators to comprehend it in its native form. So, a well-architected solution will also employ a wide range of analytics capabilities to rapidly identify and call out the most critical information collected in the data set. There are many different types of analytics designed to provide insights into collected data. Below is a list of the types of analytics most effectively employed for security intelligence, though others may apply as well:

- **Descriptive Analytics** – summarizes (or describes) events and conditions by condensing excessive information down to just the most relevant details
- **Quantitative Analytics** – relates contextual states and conditions as numerical quantities and statistics
- **Prescriptive Analytics** – employs modeling, data mining, and machine learning techniques to determine historical trends and provide guidance on optimal improvements
- **Diagnostic Analytics** – identifies the root cause of failure events and incidents of performance degradation
- **Predictive Analytics** – interpolates historical trending information to determine future states of an environment
- **Exploratory Analytics** – helps identify useful information in a data set when specific points of investigation are not predetermined (such as by looking for unusual correlations or anomalies in statistical data)
- **Inferential Analytics** – provides conclusions, states, and conditions that are not readily observable from reviewing raw data sets
- **Causal Analytics** – discovers how changes in environment elements affected outcomes in the past in order to determine the causal relations between infrastructure components

Native capabilities of a storage intelligence platform can be extended through direct integration with third-party management products. Direct integration implies a solution set is able to leverage collected data, reports, and/or automation from third-party platforms without needing to perform any additional coding or configuration. Storage intelligence platforms should establish direct integrations with other storage management solutions as well as systems, networking, virtualization, and backup and disaster recovery platforms. Additionally, restful APIs should be provided to enable the easy establishment of indirect integrations with third-party platforms. Since storage performance can be affected by elements that functionally exist outside the storage ecosystem (such as unrelated network traffic, CPU performance, and data center environmental conditions), end-to-end holistic visibility is only attainable through a federated sharing of information across management solutions.

Functionality

Storage intelligence encompasses a broad range of capabilities essential to providing guidance on the reliable and responsible storage configuration necessary to support business IT service. Some product sets include unique features that perform very specialized tasks, so each organization should carefully identify and prioritize the capabilities most applicable to its business requirements before initiating a product comparison. However, for the purpose of this evaluation, EMA identified several specific areas of support that storage intelligence solutions should include in order to be considered a comprehensive platform. Key points of functional comparison in EMA's evaluation include:

Inventory and Asset Management

Before understanding the state of the storage ecosystem, it is essential to identify and record all hardware and software components that will be audited and monitored. This requires access to a comprehensive asset database that provides granular details on storage resources, including device models, locations, configurations, filesystem details, virtualization instances, and ownership within organizational structures. Centralized inventory and asset databases should be propagated by automated storage asset discovery processes or through integration with third-party asset management platforms, CMDBs, or listing services (such as Active Directory).

Status and Performance Data Collection

Granular detail should be collected from all managed storage devices in a support stack to provide a complete picture of the state of service performance. This includes details on SAN and NAS components, such as embedded gateways, fiber channel switches, fixed SAN switches, and director-class SAN switches. Collected status information on storage units should include current capacity details, resource availability (i.e., uptime), data density, and durability states (i.e., RAID, replication, and/or erasure coding levels). The number of possible storage-related performance indicators is far too numerous to include here; however, the broader and more granular the data collection, the more accurate and relevant results will be attained from the analytical evaluation. Below is an abbreviated list of some of the more valuable storage-related performance indicators as determined by EMA:

- Frame rate
- Packet headers
- Transmission errors
- Number of hops between switches
- Inter-switch link configurations
- NPIV status
- HBA queue depth
- Multipath status/performance
- Storage processor cache dirty pages
- Storage processor utilization
- Storage processor response time
- Storage processor port queue full count
- Utilization
- Queue length
- Average busy queue length

- Response time
- Service time
- Total throughput
- Write throughput
- Read throughput
- Total bandwidth
- Read bandwidth
- Write bandwidth
- Forced flushes per second
- Full stripe writes per second
- Cache variable consumption rate (CVCR)
- Used prefetches
- Disk crossing
- Disk crossing per second
- Read cache hits per second
- Read cache misses per second
- Reads from read cache per second
- Reads from write cache per second
- Read cache hit ratio
- Write cache hits per second
- Write cache misses per second
- Write cache rehits per second
- Write cache hit ration
- Write cache rehit ratio

Capacity Management

Data collected on storage capacities should provide a complete picture of usage rates so proactive process can be introduced to ensure continuous space availability. At minimum, this requires a historical trending on available capacity. Ideally, analytics will be applied to trends to determine and alert when existing capacity will be exceeded. The addition or integration of automation capabilities can also be leveraged to automatically expand capacities if/when established thresholds are exceeded (for example, through cloudbursting). Additionally, capacity management features should identify opportunities to reclaim and repurpose unused allocated storage.

Virtualization Support

Virtualization technologies can mask storage-related issues and the root cause of problems. Storage intelligence solutions, therefore, should have the ability to look across virtualization layers and map VMs to associated storage infrastructure. Additionally, since performance of virtualization servers can have a direct impact on storage availability, status information should be collected and presented on relevant virtualization servers.

File and Content Management

Not all data has the same inherent value. Some files and data records necessitate more stringent security and superior high-availability in order to meet regulatory and business requirements. Storage intelligence solutions should be able to automatically distinguish types of data records and their value by mining metadata or content. Data characteristics that are particularly helpful for determining content value include file owner, creator, frequency of access, type of content, and level of content confidentiality.

Workload Modeling

Collected data should be interpolated to automatically create digital models of storage systems and workloads. These models may represent a particular snapshot in time or indicate current conditions in real time. Predefined workload models should be included with the platform that can be customized to meet each organization's unique requirements and configurations. An ideal solution will leverage analytics to create hypothetical models based on additions and changes to workloads. In this way, organizations can proactively identify what effects changes to the environment will have on production services before implementing them.

Report Generation

Out-of-the-box storage intelligence solutions should provide a wide range of prebuilt reports designed to provide insights into the most common storage information requirements. These should include historical trending and status reports indicating response times, IOPS over time, capacity, fault detection, and workload performance. Reports should be customizable so they can be easily adapted to meet each organization's unique requirements. Additionally, facilities should be provided to easily generate custom reports and dashboards. Custom report creation is dramatically simplified with a graphical interface, selectable fields and formats, drag-and-drop building tools, Boolean logic rules, and an integrated scripting interface.

Backup and Disaster Recovery (BDR)

Storage intelligence solutions should support business continuity by reporting on the success rate of backup services and the status of backup systems. This includes identification of backup system capacities and confirmation on the integrity of backup and archive data. Alerts should be sent on detection of backup failures or any conditions that could inhibit disaster recovery processes.

Financial Analysis

Understanding the financial impact of storage resources and services is essential to introducing more cost-effective implementations that align with business requirements and budget. Storage intelligence platforms should identify and report on storage-related infrastructure costs, software costs, maintenance costs, managed service costs, and cloud service costs. This information can also be used to provide chargeback or billing reports to individual departments or customers.

Deployment and Administration

The ease with which a solution can be deployed is directly related to the complexity of the infrastructure supporting it. The more hardware and software elements need to be installed, the more challenging the deployment will likely be. An ideal solution will employ automation for enabling a turnkey deployment process, rapidly installing software components (such as databases, reporting engines, and console interfaces) and automatically detecting the mobile endpoints that will be supported. If data collection agents need to be deployed on managed devices, they should be automatically pushed from the console server rather than requiring manual installations. Cloud-hosted solutions have an advantage in deployment processes as they typically do not require any additional hardware or software installations prior to activation. Similarly, appliance-based solutions typically require few installation tasks other than physically connecting them to a network and powering them on.

Administration practices are simplified with the use of an intuitive and customizable console interface that incorporates all the storage intelligence processes, dashboards, and reports into a single pane of glass. Consoles built with an HTML 5 interface enable administrators to perform management tasks from any network-connected devices (including PCs and mobile devices). The processes for collecting storage asset and status data should be automated, requiring little or no administrator interaction. Role-based profiles that are integrated with listing services (such as Active Directory) also simplify administration by allowing users to be organized into logical groups (e.g., by job function or department) and be collectively supported with a common set of permissions and restrictions.

Vendors must also display a commitment to supporting the storage intelligence platform and their customer communities. Maintenance contracts should be offered that deliver responsive and continuously available live support, as well as timely product updates. Vendors should offer professional services staffed with support professionals who are knowledgeable about their solution set to assist customers with

training, problem solving, environment optimization, and the initial product deployment. Vendors should also engage the user community by hosting online forums and regular meetings to educate organizations on the effective use of their platforms and storage management best practices.

Cost Advantage

Pricing models for a storage intelligence platform should be simple to understand and easy to calculate. The platforms EMA evaluated for the “EMA RADAR for Storage Intelligence” fall into two distinct categories: solutions designed to support vendor-specific storage platforms, and independent software solutions designed to support heterogeneous storage ecosystems. Vendor-specific solutions are typically offered with the purchase of the vendor’s storage platform, along with an accompanying maintenance contract for no additional cost. Independent solutions were most commonly offered for per-gigabyte licenses sold for either a one-time cost or annual subscription fees. Vendor-specific solutions and platforms licensed on a subscription basis typically did not require any additional cost for maintenance contracts; however, one-time perpetual license customers often must pay an additional annual maintenance fee (usually calculated as a percentage of the total license cost) to gain access to platform updates and the vendor’s help desk support. When choosing between perpetual and subscription licensing, it is helpful to recognize that the breaking point is between two to three years—in other words, after that amount of time, accumulated subscription costs exceed a one-time purchase cost, including maintenance costs. In deference to the reality that most organizations retain management solutions for many years or even decades, EMA’s evaluation models reviewed total costs of ownership over three- and seven-year periods. As a consequence, the evaluation results did favor perpetual license solutions.

Vendors may offer on-premises solutions, cloud-hosted solutions, or both. While cloud-hosted solutions are most commonly offered for subscription pricing, on-premises solutions may be offered for either subscription or perpetual licenses. Also, while cloud-hosted solutions require no upfront costs or pre-deployment efforts (other than service registration), on-premises solutions require the purchase and installation of a physical server or appliance, its operating system, and possibly an SQL database. EMA recommends enterprises perform price comparisons that take into consideration all cost elements (license, maintenance, and infrastructure costs) and recognize their organization’s unique requirements. Purchasing a low-cost or vendor-specific solution may impede an organization from achieving long-term storage management goals by limiting visibility. Purchasing a more comprehensive solution may not be cost-effective if the advanced features are never used. It is important to right-size a solution based on projected future requirements as well as existing goals.

Vendor Strength

Consumers should always be aware of a vendor’s stability and its commitment to a platform prior to adoption of the solution in order to be sure of its long-term viability. A vendor that is financially strong with high revenue and vast equity is more likely to continue support for a management platform. Solution providers that invest heavily in research and development will also be assured of maintaining continual value in the platform’s architecture and feature set. Strategic and channel partnerships also increase vendor relevance in the market space and customer loyalty extends visible credibility. Additionally, a vendor’s vision and strategy for development, innovation, and foresight of future requirements indicates whether a management solution will maintain optimal value in a dynamic marketplace.

Evaluation Criteria

Feature Eligibility

In order for a product set to be credited with a feature or capability in EMA’s evaluation, it was required to meet three strict criteria:

- The features needed to be generally available with the solution set at the time of the evaluation. Any features that were in beta testing or were scheduled to be included in later releases of the management suite were not eligible for consideration.
- All features needed to be self-contained within the included package sets. Any features not natively included in the evaluated package sets but available separately from the same vendor or third-party vendors for an additional cost did not qualify.
- All reported features needed to be clearly documented in publically-available resources (such as user manuals or technical papers) for verification.

Financial Evaluation

To enable product license cost comparisons that are as fair as can possibly be attained through analytical process, EMA developed six sample support models and applied vendor pricing to each. Pricing included license or subscription costs for all products, add-ons, and modules necessary to achieve the functionality credited in all other sections of this evaluation. Additionally, expenditures were added to account for any additional hardware and/or software infrastructure costs necessary for the platform to operate, and maintenance costs for the platform were calculated for the time period specified in each model. The results for each of the six models were empirically rated on a pricing scale (i.e., rated from 1-10 with a two decimal point level of accuracy). Ratings for all six models were then averaged to provide the final scoring reported in this evaluation. The six models used in EMA's evaluation are shown here:

- Short-Term Small Business Model – supporting 100 users (140 devices) over 3 years
- Long-Term Small Business Model – supporting 100 users (140 devices) over 7 years
- Short-Term Medium Business Model – supporting 1,000 users (1,400 devices) over 3 years
- Long-Term Medium Business Model – supporting 1,000 users (1,400 devices) over 7 years
- Short-Term Large Business Model – supporting 10,000 users (14,000 devices) over 3 years
- Long-Term Large Business Model – supporting 10,000 users (14,000 devices) over 7 years

Organizations that offer multiple product licensing and/or deployment models were evaluated and scored across all potential scenarios, and the best scores achieved were included in the final review. For vendor-specific platforms included with the purchase of a storage platform, costs of the storage platforms were sized to meet specifications for each model and maintenance contracts were calculated. Since maintenance contracts are required in order to gain access to these solution sets, these annual fees were determined to be the cost of the storage intelligence solutions for the purposes of providing side-by-side costs comparisons in this report. However, it is important to note that organizations may need to purchase these maintenance contracts to support their storage hardware infrastructure and may consider the availability of the storage intelligence platform as a free add-on service. EMA urges organizations to perform their own cost-comparison based on internal business requirements and budget considerations.

On the EMA RADAR

Storage Intelligence Market Overview

Value in any solution can be clearly defined by comparing the strength of the platform with its cost effectiveness. The EMA Storage Intelligence Market Landscape Chart below provides graphical representations of evaluated industry leader positioning in relation to both critical axes. The “Product Strength” axis combines evaluation scores for *Functionality* with *Architecture & Integration*. “Cost Efficiency” is calculated by adding the scores achieved for *Cost Advantage* and *Deployment & Administration*. The size of each bubble indicates scoring for the *Vendor Strength*.

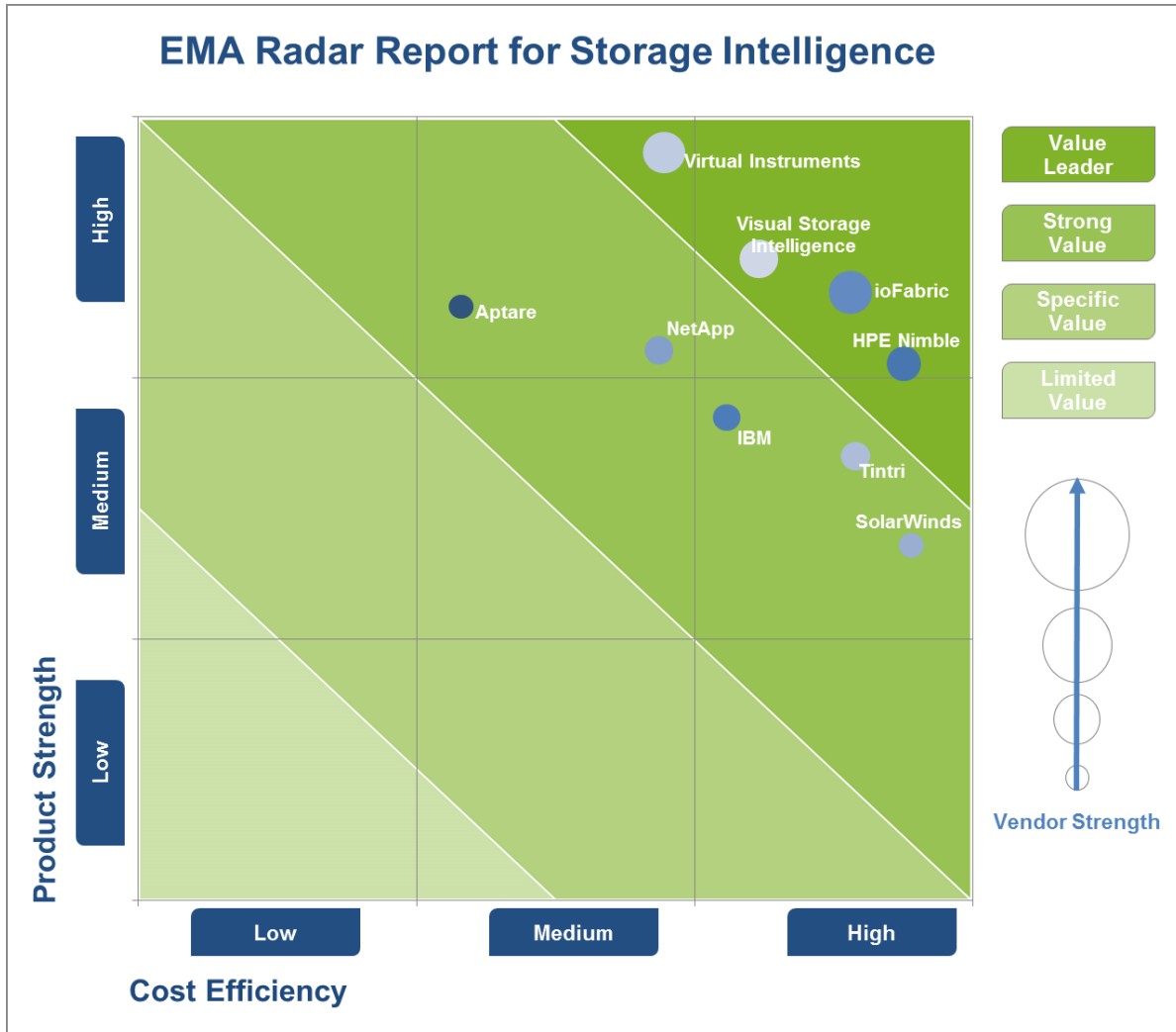


Figure 1: EMA Storage Intelligence Market Landscape Chart