White Paper

Real-time IT Visibility: Considerations for Buyers

Questions to Ask when Choosing a Solution

By Dan Conde and Nik Rouda, ESG Senior Analysts

August 2016

This ESG White Paper was commissioned by ExtraHop and is distributed under license from ESG.

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

Contents

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

Introduction

Information technology is at the core of the modern enterprise. Nearly every facet of an organization relies in some way on technology, whether it's connected manufacturing devices, enterprise resource planning, marketing automation, or financial planning. At the same time, the infrastructure that supports these resources has grown more complex. Disruptive technology trends like cloud, SDN, IoT, and virtualization are changing the landscape of IT at a time when end-users rely on IT more than ever before. It is this end-user reliance, coupled with scale and dynamism, that calls for a new approach to IT monitoring. Modern monitoring must be real-time, at scale, and capable of providing visibility wherever IT assets reside. IT operations analytics is at the forefront of this monitoring evolution, but understanding how these often diverse technologies fit together, and what benefits they actually deliver to IT, can be complex.

ESG would like to offer a framework for evaluating IT Operations Analytics (ITOA) technologies, including key questions to ask prospective vendors around deployment and management, depth and breadth of visibility, scale, data quality, and extensibility.

Monitoring Needs

A common business trope is that "you can't manage what you can't measure," pointing to the need for visibility into operations to influence strategy and decision-making. There is certainly wisdom in this concept, and it's surprisingly applicable to the modern enterprise IT environment. As vendor offerings become more sophisticated, complexity continues to rise, and many CIOs and IT directors find themselves losing clarity about which assets are exactly where and how they are performing.

Guiding Principles

A few basic guiding principles might be helpful in finding a way out of this mess.

- Installation and configuration need to be simple.
- Data collection, aggregation, analysis, and reporting must integrate neatly with current operations and workflows.
- IT professionals need to be able to quickly self-identify and answer questions like: Do we have an issue? Are we getting complaints? Are there application errors? How do I fix this? Can I go home now?

The unfortunate result of complexity is that several IT systems have become "black boxes," which is okay when the system is working well, but a nightmare when changes need to be made or things go wrong within interdependent environments. This problem extends beyond the walls of the data center and spans all offices and mobile workers. So it's no surprise that recent ESG research revealed that nearly half of organizations with big data analytics initiatives cited improved operational efficiency as a desired business benefit from these efforts.[1] Their hope is that more advanced tools will offer increased understanding of the nuances of applications and infrastructure for the business. Indeed, the most commonly identified data sources for analytics exercises are server, storage, network, and other IT logs, demonstrating the widespread use of technology to track and manage technology.[2]

In short, modern IT teams need solutions that allow them to efficiently identify, monitor, and derive insights across the environment. Any monitoring solutions should be evaluated first and foremost on these pillars, and with that in mind, we've created a framework to help buyers evaluate and rationalize the myriad choices on the market.

[1] Source: ESG Research Report, *Enterprise Big Data, BI, and Analytics Trends: Redux*, July 2016.

[2] ibid.

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

A Buyer's Framework for Gaining Access to IT and Business Data in a Dynamic IT Environment

In a dynamic environment, the ability to access the right data is the foundation of a strong monitoring solution. Incomplete visibility into the environment leaves gaps, which may lead to misinterpreting critical information. ESG has broken down this framework to five key categories of deployment and management, depth and breadth of visibility, scale, data quality, and extensibility to help you choose a holistic and robust monitoring solution.

Deployment & Management

Instrumentation and monitoring need to be non-invasive and quickly deployable and must not introduce new security risks. If the task of monitoring starts to degrade the effective operation of systems, it defeats the purpose, and creates internal resistance in deployment, which reduces visibility. Organizations must have the ability to rapidly classify IT assets, discern their purpose, map their relationships, and identify their dependencies. All assets need to be understood in context with appropriate consideration for how they affect IT operations. For example, a mission-critical device should be categorized differently than a pilot or test-phase asset.

Consideration Questions:

- How complex is the deployment process?

  If the sales engineer says, "It installs in just 15 minutes," you should walk them out. There is far more to it than just installing the bits. Make sure to account for three major things: 1) the number of touchpoints required, such as agents, probes, sensors, collection points, network taps, etc., 2) the amount of infrastructure needed to deploy the solution, such as rack space, power, servers, etc., 3) the configuration & management required before value is realized, such as setting up analytics engines, dashboards, reporting, *and* having the ability to gather and process enough data to deliver a complete view of your environment.

  For example, solutions that don't collect server data may not have server agents or probes, which could save some effort in installation on 10,000 nodes for sure, but may limit access to info you need and in turn, limit your visibility. On the other hand, who wants to spend weeks manually loading and configuring software on each server, device, router, switch, etc.? No one, because the time and cost of managing this type of deployment is infeasible. Further, if you want to understand an event that happens monthly, like a push of antivirus updates, you are looking at a month or more until a pattern emerges from which you can derive any actionable insight.

- What is the support structure around the platform?

  A solution isn't just composed of hardware and software. To be effective in an enterprise, any solution should have the proper support structures that assure success: professional installation and operationalization, documentation, training, and an accessible community for help. Yes, quality technical support is part of the picture, as are professional services to assist with projects when necessary, but having a richer set of resources readily available can make a huge difference in outcomes. The total cost of ownership depends on how easily the platform can be run; no one wants to waste time troubleshooting a troubleshooting tool.

Scale

Scalability is another key consideration, which has several important dimensions. Data acquisition must be done at high volumes, across many sources and types of data formats. Also important is the number of concurrent users and the ability to support the analytics workload demands they place on the system.

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

Consideration Questions:

- How scalable is the platform in a dynamic heterogeneous environment?

  Many tools appear to perform well in small proof-of-concept lab tests, yet fall down when applied to a large enterprise with many simultaneous users. Being able to drill deep into large volumes of traffic, across thousands of nodes, applications, users, and many geographic locations, is critical. A lot of variety can be found in today's heterogeneous environments, such as the types of networks from LANs, different service providers' WANs, and VPNs, to operating systems such as Linux and Microsoft Windows. It is well worth exploring how extensible and scalable the platform really is to meet the volume and variety of data. Sometimes this requires a more comprehensive test, but at a minimum, buyers should check references for existing customers of similar size and complexity to their own environments.

- What are the product's limiting factors when it comes to scale?

  Every product has its constraints, so make sure to thoroughly analyze what they are and how they can affect your monitoring capabilities. Is the product limited by the number of devices it can monitor at any given time, the number of current users logged into the platform, or the amount of throughput or data ingestion rates that it can handle? Can it geo-scale and/or provide insights into remote sites? Architecture can also be telling here, so ask how data is captured, stored, and processed, and what the limiting factors are in terms of memory, processor, and disk. Not every constraint applies to your situation, but make sure to weigh them accordingly when investigating a monitoring solution.

- How does the vendor define "real time"?

  Few terms get as overused as much as "real time," especially when it comes to analytics. Yet there are many possible definitions for what real time means and whether it is fast enough to meet your actual requirements. When you think of *fast* data in the context of IT infrastructure, you might be referring to fast ingest, fast preparation, fast analytics, fast diagnosis, or fast response. A vendor's definition of real time should encapsulate all of those steps for the data as they happen, and better lead to an immediate course of action. Data processing performance also matters because, otherwise, you will have to reduce the amount of data you can collect and understand. Not least, total response time for issues that are impacting business operations is a critical component of service level agreements. A real-time solution should consider all aspects of the problem and response. Buyers should verify which elements are actually delivered in real time before purchasing a solution.

Depth & Breadth of Data Visibility

To be effective, monitoring solutions require breadth of visibility across the application delivery-chain and the organization, providing a common viewpoint of the data. Appropriate solutions need to present enterprise-wide visibility from on-premises, the cloud, and remote locations along with mobile and the growing use of IoT devices.

Consideration Questions:

- What data sources can be utilized (and with what limitations)?

  Understanding your environment starts with the ability to access *all* relevant data, so it's critical to understand a product's data source and the visibility it affords. Many will suggest that a single data source is adequate, such as log files or network packet capture, but this is blatantly untrue. No data source covers every perspective;

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

however, some are much more robust and expansive than others. For instance, code-instrumented agents provide visibility from within the application (threads, memory allocation, etc.), and logs provide self-reported machine data that might not traverse the network (services, errors), while network data provides passive visibility of everything communicating over the network. For example, an application could make a database query and get no response. The agent would just know a query was attempted. The machine data would show an error if logging levels were configured adequately. The network data could determine if it was a network or database issue but not which service initiated the request. Each part of the puzzle is critical, but incomplete without the other data sources. Correlating activity across as many data sources as possible gives the richest understanding. Savvy buyers should also go beyond a checkbox list of sources to see exactly how each vendor acquires data and handles it. Again, as an example, if a product does packet capture, how long does it keep that information? Where is it held? Is it secure there (it surely contains sensitive info)? What are the long-term costs of storing all that data? Asking follow-up questions that show how data sources are collected, kept, and leveraged will bring to light important differences between products.

- What's the depth of application visibility and the breadth of infrastructure or resource dependencies?

  Every application has nuances in how it's architected, deployed, behaves, and communicates. Being able to understand or even completely reconstruct the application's behavior and communication patterns can be extremely useful to diagnose a problem or to audit activity. Yet many applications make monitoring non-trivial, requiring agents or specialized utilities and sometimes even requiring monitoring tools to be trusted hosts. As examples, Skype can change ports to get around blocked access, while Microsoft Exchange can encrypt communications. Following these common applications and correlating them with the downstream resources they utilize is critical to application-awareness and should be considered when providing visibility for these environments.

- How applicable is the platform across different groups within the organization?

  There are many disciplines within an IT department, and rarely can they function completely independently. Most often, different specializations must collaborate in order to understand the full picture. However, if a monitoring platform provides information and insights to particular groups or must be decoded by highly skilled specialists, collaboration can become difficult, if not impossible.

  A robust platform should be accessible and meaningful to people who view issues from different perspectives—not just networking or virtualization but servers, storage, security, applications, and cloud. Bringing these groups together onto a common platform allows a holistic comparison of all aspects and implications of a significant trend or major incident.

Data Quality

There are three primary considerations surrounding data quality. First is the source of the data itself, whether that is logs, agents, probes, or the network. Second is how the data is gathered, as the collection itself can impact how easily and readily the data can be analyzed. One key area of focus around data gathering is continuous and automated discovery of IT assets (e.g., infrastructure, applications, devices, etc.) Enterprises are dynamic and a map of the infrastructure (and thus, data inputs) changes constantly. The third consideration is how easy it is to correlate that data with other sources. No one data source can deliver insight into everything, so the context provided by other data sets is critical to ensuring quality.

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

Consideration Questions:

- How comprehensive and/or rich is the data set?

  Often, the data can be superficial, requiring IT to drill down to get the details to provide context and insights. The data needs to span tiers, starting at the application level and descending into the operating system, and finally into the hardware infrastructure that includes the network and storage. The solution must be supported for many users working with very large data sets generated by many hosts, all at the same time.

  Architectures vary in the IT monitoring space, and this variance produces disparate results. Sampling data may be desired in order to reduce the data footprint and keep costs down, but may also lead to missing important data points. Outliers, rare activity, and subtle patterns might be missed. Conversely, saving every packet or log file forever would surely make your storage sales rep very happy, but probably not your CFO. The amount of overhead is costly. How and when data is ingested has an impact on how well the system can keep up with volume and velocity; certainly, it would be unacceptable to delay a transaction just to record it in a debug log. It's important to find a product that balances the visibility required with the resources needed to provide that visibility; however, you should strive to not sacrifice any performance, accuracy, or effectiveness.

- (Optional) How trustworthy is the data? (Is it empirical or self-reported? This question should also be asked with all other questions.)

  Trustworthiness of the data can affect the quality of the data, and can overwhelm other considerations such as how the data is gathered.

- (Optional) What is the time to value of the data? (data scientist vs. turnkey analytics)

  The value of the data often depends on the freshness of the data–thus you should understand who will use the data and when they will need to use it.

Extensibility

Modern IT is rarely a homogenous island. Integration and interoperability are key to any solution deployed into the environment. Integration with additional data types and sources, such as logs, agents, and network (including wire data) are essential, not to mention machine learning techniques and visualization tools. The solution must also look towards the future and adapt to your dynamic IT environment without major software releases or costly upgrades.

Consideration Questions:

- Does your platform integrate with other solutions?

  For any visibility solution, your insights will only be as useful as their applicability across your environment. Unless your business was started yesterday, it is very likely that you already have a few management tools in house. If your monitoring solution doesn't integrate with these existing software, processes, and workflows, it will be limited in terms of its scope of use and hurt operationalization of the technology. Ask yourself: is it an open architecture? Is the data portable? Is there support for APIs and a bi-directional flow of data?

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

Data Analysis

Once data is acquired, it must be analyzed in a manner that is easily applied to derive insights. Here are some data analysis capabilities to consider.

Table 1. Data Analysis Capabilities Considerations

Data Analysis Capability	Considerations
Ability to respond to changing conditions	The context in which the data is monitored and used changes frequently. Therefore, a system must have access to the data in real time, and not show out-of-date archives, to keep up with dynamic environments. The data must be consistent and reliable, and the system must provide agile management capabilities and conditional alerting.
Ease of use	The solution ought to have a unified UI to provide consistency across usage areas, and intuitive data analytics to enable non-experts to interpret the data. If a tool is kept in the hands of a few experts, it will not provide the full potential value.
Ability to analyze data across tiers	Data correlation capabilities are needed to understand data that may be gathered at the networking layer but may affect functionality at the application layer.

Practicality

These systems must be practical and easily consumable in an existing IT environment workflow. Complexity in acquisition, deployment, and operational capabilities leads to solutions that are not used where they are needed the most.

Table 2. Practical Considerations

Characteristics	Considerations
Cost	Costs and vendor licensing models must be flexible enough to meet environmental requirements as they change, and should not force buyers to make unnecessary purchases or introduce delays.
Learning curve	The barrier to adoption must be suitably low and allow users to implement the monitoring solution within the existing confines of their workflows. This reduces the time to value and ensures the solution is operationalized.
Managing at scale	The system itself shouldn't add incredible complexity to the environment or impose severe management burdens. A monitoring solution that requires excessive management to operate won't be used to its full potential.
Scale with the business	As the IT environment grows, the monitoring solution must also be able to grow. There is a need for scale related to the amount of machines it can handle, the raw amount of data it can process in real time, and the new technologies that it can add into the mix.
Security impact	Systems should not negatively impact an organization's security posture. It would be truly tragic if a security monitoring solution created new vulnerabilities.

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

The Bigger Truth

Organizations considering different solutions for improving IT visibility should not fall into the trap of simply performing a one-by-one comparison of features or confirming that the features conform to a check-off list. Consider how well the solutions support features and how the individual features are combined to create an integrated solution.

The choice also shouldn't be based solely on performance; while it is important, a solution's full capability will be wasted if it is not easy to operate or key features are not readily discoverable in the user interface.

More importantly, one must realize that IT environments are increasingly dynamic. Deployment types are changing, so the problem is no longer limited to the data center. Thus monitoring solutions need to strive to provide key capabilities, and also be dependable, adaptable, and fluid. An undependable monitoring solution is almost as useless as having no monitoring solution; in other words, "absence of evidence is not evidence of absence."

By choosing a monitoring solution that provides good visibility across the changing IT asset landscape, offers depth of information on each element, and is easy to adopt, IT will realize ROI and time to value that meets the enterprise's needs for improving service and security levels.

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of The Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at 508.482.0188.

Enterprise Strategy Group is an IT analyst, research, validation, and strategy firm that provides actionable insight and intelligence to the global IT community.

© 2016 by The Enterprise Strategy Group, Inc. All Rights Reserved.

 www.esg-global.com

 contact@esg-global.com

 P. 508.482.0188