

## BIG DATA APPS — MAKING SENSE OF MULTI-STRUCTURED LOG DATA

The high level of attention being paid to “big data” reflects the great potential that lies in large repositories of unfiltered machine (or log) data. One-dimensional tools to catalog and search such data fill a certain need, but companies that want to extract greater value from their machine data need multi-dimensional applications that turn log files into actionable insights that can have a tangible impact on the way a business functions.

### INTRODUCTION: THE PROMISE OF MACHINE DATA

“Big data” comes in many forms, and the tools to manage and analyze it have a range of capabilities. A particularly important form of big data is product machine data -- information about how a product or system is configured, how it is used, and how well it is performing. (By product or system, we mean any intelligent hardware and/or software device, or collection of devices, that generates and reports digital information about itself.)

Businesses can benefit from a deeper examination of machine data in a number of ways -- reactive diagnostics, proactive problem identification, and business intelligence analysis. One of the key reasons why machine data holds such potential is that it contains the unfiltered truth about products and how they are being used. Machine data represents facts, not guesses or hunches.

Unlocking the value in unfiltered machine data can only be achieved, however, by using the right tools that not only index machine data but extract important insights about system performance and user/machine behavior that can be acted upon in a timely fashion.

### ALTERNATIVE APPROACHES TO MINING MACHINE DATA

Making sense out of large repositories of machine data is a complex undertaking. The first step in mastering the task is to collect and index vast volumes of log files. Several tools exist for this purpose, and companies are adopting these tools in large numbers -- a testimony to the desire to explore what's hidden in all the big data.

First-level tools are designed to be broadly applicable and simple to install -- both of which are desirable attributes. But these qualities make these tools a mile wide and a foot deep. They can have several drawbacks.

- They usually only offer indexing and searching, leaving it up to the end user to know what they are looking for.
- They are tools for IT professionals; getting to the next level of analysis and value requires specialized skills and/or funding requests for IT projects.
- They are not integrated into existing processes such as a help desk or customer relationship management (CRM) system.
- They are not designed with the business user in mind, but are complex to learn and use.

In short, these first-level tools are valuable as a starting point for mining repositories of operational data. But they start from the perspective of a device and the log data it generates, and thus do not deliver answers to important questions posed by personnel in key business functions such as support, engineering and sales. For these users, an alternative approach is required -- an approach that begins with the perspective of the business user of operational data, not the IT department.

This alternative approach does not replace general-purpose indexing and search tools, but builds upon them in the same way that business-intelligence applications build upon general-purpose databases. Instead of saying, "How can we make it easy for IT professionals to search through log files?" it says, "How can a business professional do his or her job better through intelligent, focused analysis of relevant machine data?"

### GLASSBEAM'S BUSINESS-CENTRIC APPROACH TO MACHINE DATA

Glassbeam is pioneering a business-centric approach to analyzing machine data. Glassbeam starts with the premise that machine data contain critical business information that, when mined, provide the business with a competitive edge. To unlock this hidden value, however, a new class of tools -- analytical applications for machine data -- is needed. Glassbeam has developed such applications, based on its breakthrough SPL language (see bottom bar) that is specifically designed to extract the value of "multi-dimensional" machine data.

Log data contain not only tactical information on errors and warnings but also strategic information on subjects such as product usage, configuration, licenses, and the features enabled or used by customers. While traditional log processing tools focus on tactical events and errors, Glassbeam's multi-dimensional approach provides a comprehensive view of all data contained within all log files produced by intelligent devices. Hence, the value is not just to end user IT departments but to business users inside product companies as well.

Typical log files are semi-structured, with one format such as time series data. Glassbeam understands that log bundles often contain many types of files, each with multiple sections with time series, configuration, statistics and snapshot data and each in their own format. This is the norm for many product companies. Parsing and making sense of the data require more sophisticated tools than just search and indexing. This is where Glassbeam's applications excel.

### GLASSBEAM'S BREAKTHROUGH TECHNOLOGY

Glassbeam's ability to quickly turn vast amounts of machine data into actionable intelligence for business professionals is made possible by its breakthrough applications enabled by an advanced technical platform consisting of parsers and indexers that simplify the task of specifying, parsing and indexing rules for multi-structured data.

The platform includes a breakthrough language, SPL™ (Semiotic Parsing Language), that specifies the rules for parsing multi-structured data. Using SPL™, a business analyst can describe the rules and semantics for parsing a set of log files. For semi-structured data such as machine data logs, this means being able to take advantage of the structure inherent in the data without any pre-defined grammars. No specific understanding of databases or procedural programming is required.

The Glassbeam platform also employs advanced data warehousing technology to produce applications that are fast and can scale to handle terabytes of data.

Glassbeam's understanding of the value of multi-dimensional machine data allows it to tailor results to its customers' specific needs. The process begins by targeting the function(s) that a business wants to improve, such as customer support, new product development, sales or product management. With the target business functions selected, Glassbeam and the customer identify the specific pieces of machine data that they wish to track. A product manager or product design engineer will have interests in machine data that are very different from those of a product support specialist or the account manager assigned to a customer. So, the ability to collect, parse and analyze the specific data points of interest to each one of them makes Glassbeam's solution highly valuable.

Glassbeam's applications provide personnel in support, engineering, product management, IT and executive management with easy-to-understand dashboards that present actionable insights gleaned from machine data. These Web-based dashboards provide dramatically improved visibility into how customers are configuring and using their products, as well as how those products are performing in the field.

Critical to this business-centric approach is the realization that multi-dimensional log files can contain many different types of information, which can be of value to different functions within an organization. Let's look at some of these different types of information and how they might be used.

- **System identification-** such as serial number(s) for elements in the system. This information is essential to matching the source of log data with a specific customer.

- **Configuration information-** such as software version, model number, hardware configuration, manufacturer information or file system configuration. Changes in software version or subsystem configuration can have a huge impact on performance. Knowing which customers have specific versions can help establish patterns in key metrics, such as performance problems or part failures.

- **Statistical data-** such as response times, throughput, disk I/O, network errors. These are the core pieces of data that provide insight into how customers are using a product and how it is performing.

- **System messages-** such as events and alerts, frequency of events, levels of severity or urgency. In addition to providing real-time notice regarding something that requires attention, a collective knowledge base of these messages can contribute to trend analysis, forecasting and the design of next-generation products.

## BENEFITTING FROM A BUSINESS-CENTRIC APPROACH

A rich set of unfiltered machine data provides a broad context for extracting meaningful insights for practical business purposes. Some examples include:

- **Customer support.** Operational data on issues such as product performance, faults, errors and other anomalies can help support personnel identify the sources of problems more quickly or in some cases even proactively spot changes that need to be made to avert problems for customers.

- **Product engineering and product management.** Engineers and marketers charged with designing the next version of a product traditionally rely on subjective means of assessing customer feedback and requests. But with properly designed analysis of machine data, product engineers and marketers can base their decisions on hard facts from large numbers of users, not qualitative guesses.



■ **Sales and professional services.** Machine data can provide detailed and up-to-the-minute information on key metrics, such as amount of disk space used, number of site licenses, and software version, that can alert sales people to opportunities for new revenue. Similarly, professional services staff can turn the results of machine data analysis into a service they can sell through to their customers. (For example, a manufacturer of data-storage devices can help their customers monitor and optimize their storage systems.)

■ **Senior management.** Executives who want a quick look at how their products are performing with customers can call up an operational dashboard with easy-to-understand graphs and charts that reflect comprehensive, real-time, unfiltered facts.

Businesses can see tangible benefits and quick ROI from Glassbeam's operational data applications, among them:

- Reducing support costs by optimizing troubleshooting and case resolution.
- Increasing revenues through new value-added services for their customers.
- Providing upsell and cross-sell opportunities for sales managers.
- Increasing sales bandwidth to focus on new sales opportunities vs. spending time on data collection for account management tasks.
- Gaining access to unfiltered truth on product usage from the installed base to allow them to build better products.

## CONCLUSION

Businesses can benefit significantly by moving beyond a one-dimensional, device-centric approach toward machine data to a multi-dimensional, business-centric approach. The key is applying tools capable of uncovering business intelligence contained in multi-structured log files to specific business functions.

For more information, contact us at: [sales@glassbeam.com](mailto:sales@glassbeam.com)

Glassbeam, Inc.  
530 Lakeside Drive, Suite 100  
Sunnyvale, CA 94085  
408-740-4600  
[www.glassbeam.com](http://www.glassbeam.com)