Glassbeam White Paper

Machine Data Analytics — Solution Vs. Tool Approach

Contents

Introduction

High-tech manufacturers create products that might be software only or a combination of hardware and software. All of these products capture product operating environment details and store them as entries in log files.

These captured log data entry details vary widely among manufacturers. Glassbeam offers manufacturers several distinct advantages over competitors when it comes to delivering a comprehensive solution for machine data analytics.

Eight Key Areas of Glassbeam Differentiation

1. Data Complexity

The data structures for log files can vary greatly in form among manufacturers — and even within a single manufacturer's product suite. Glassbeam is unique in its capability to process multi-structured files including:

- Highly structured log files; for example, JSON, XML, and CSV
- Structured log files; for example, time-stamped log entries with common text format
- Semi-structured log files; for example, log entries including time-stamped entries in a mixed-text format
- Unstructured log files; for example, log entries that may or may not include time-stamp entries and are typically in a mixed text format
- Email; customer support email is a common input to the customer support organization

2. Data Streams

Streamed data is inherent to Glassbeam's architecture. Glassbeam processes all data as if it were streamed, regardless of whether the source data received was a bulk upload or a real-time data stream. The entire Glassbeam system platform is built on the premise that data is processed as soon as it becomes available – NOT processed as batches. Acting upon data as soon as it becomes available is a critical requirement for many use cases, including near real-time and real-time analytics demanded by IoT (Internet of Things).

3. Purpose-built Cloud-based Architecture

Glassbeam machine data analytics is a platform consisting of a number of applications tailored to perform specific functions and coordinated to ingest, process, and transform multi-format log files into consumable, structured data formats. These structured data formats support data analysis and indexed search queries through our InfoServer middleware, providing end-user access via RESTful APIs.

4. Rules & Alerts

Glassbeam extends the value of log analytics use cases by applying rules. The Rules & Alerts Engine and its FSM (Finite State Machine) can be applied to further process source data supporting use cases to extract additional elements of contextual information. The complex rules defined enable proactive action based on a specific circumstance or set of circumstances.

5. Machine Learning

Integrating the Rules & Alerts Engine with MLlib (Machine Learning library) enhances Glassbeam's current machine learning capabilities. MLlib is Apache Spark's scalable machine-learning library consisting of common learning algorithms and utilities including classification, regression, clustering, collaborative filtering, dimensionality reduction, and optimization primitives. Use cases, plus the application of the Rules & Alerts Engine and MLlib, advance the role of analytics into the realms of machine learning with predictive and prescriptive analytics.

6. Domain Expertise

Glassbeam's current customers are all product manufacturers. Glassbeam has gained tremendous experience and deep knowledge around key business functions of support, services, engineering, sales and marketing in the product manufacturer organization.

7. End-to-end Solution Delivery

A Glassbeam customer doesn't visit our web page to download a log analytics tool. Glassbeam offers its customers a machine-data analytics Platform designed to deliver an end-to-end solution. An end-to-end solution begins with discovery sessions between customer participants and Glassbeam Professional Service Consultants to understand relevant details encompassing the customer's business and technical domains.

Once the domain target for analytics is understood, Glassbeam works with customers to discuss standard use cases and to define additional customer-specific use cases to deliver the analytical results required. The next step in the solution delivery model is ingesting, transforming, and loading the transformed results of multi-structured log files into up to three types of data stores. Analytical results are presented to customer users via a combination of ad hoc and purpose-built dashboards, supplemented with the added capability of performing indexed search queries for analyzing log files for details falling outside the scope of the current set of dashboards.

8. Total Cost of Ownership

As mentioned earlier, Glassbeam offers an analytics solution that includes Glassbeam subject matter expertise, a platform for ingesting multi-structured log files, and output of rich analytics to users in the form of dashboards and indexed search queries. Thus, Glassbeam concentrates on delivering superior Total Cost of Ownership (TCO) over price per gigabyte (GB). Based on the specific customer circumstances, TCO incurred by our customer is either minimized or negated by the Return on Investment (ROI). Glassbeam customers have accrued impressive ROIs by reducing support costs, improving customer satisfaction, and increasing revenue. Examples include:

- Fortune 50 Technology firm: achieved an ROI of $7.5 million in headcount savings of 30 TSRs and new upsell opportunities.
- Fortune 500 Medical Devices firm: achieved ROI of $3 million by a 90% reduction in time to to aggregate and searching through logs.
- SMB Technology leader: achieved ROI greater than 25% from efficiencies in support processes and centralization of tribal knowledge.

Comparing Glassbeam's Solution Technical Approach

Glassbeam competitors include DIY (Do It Yourself/Build Your Own) analytics and log analysis application tools like Splunk, SumoLogic, and Loggly. The focus of this paper is comparing Glassbeam's platform solutions approach to machine data analytics to analytics applications employing an indexed search tool approach.

There is no question that vendors of applications promoting the indexed search approach to log analysis have achieved great market success in promoting the tool

value of their analytics. The success these vendors experienced began with their growth inside the enterprise IT environment where there was a need to analyze well-known log file formats (like syslog files) common to this environment. The enterprise IT environment continues to be the primary sweet-spot for the continued growth and success of these log file analysis applications.

Glassbeam acknowledges the strong fit of the indexed search tool approach to log file analytics for the enterprise IT department and other departments that have well-known log file formats. However, for those enterprise departments having semi-structured and unstructured log files, Glassbeam boldly steps up to any challenge by these vendors through the value offered by our machine data analytics solution that we, and our customers, believe is measurably superior to the indexed search tool approach to log file analytics.

As mentioned earlier, Glassbeam takes a solutions approach, offering its customers machine data analytics for extracting usable and valuable information from log files. Unlike the indexed search tool approach to log file analysis, however, Glassbeam doesn't offer instant gratification. Indexed search tool vendors make their analytics application available as a download link on their web sites. If a user downloads and installs one of these analytics applications, and if the user's log files are acceptably structured and time-stamped in accordance with the application's expectations, the user will have the ability to view and search the log files indexed.

Conversely, Glassbeam's solution is a Machine Data Analytics platform whose focus is on the long-term satisfaction of its customers. Our solutions approach is not an application downloadable from the web. Glassbeam's analytics platform does the heavy lifting for customers by removing the burden of having to figure out how to ingest and process huge volumes of semi-structured and unstructured data formats.

Our customers do not have to worry about writing programs and creating cryptic REGEX (Regular Expression) code to transform log data into a usable format. You may recall, for the indexed search tools to provide instant gratification, the application has the expectation that data will be acceptably structured and time-

stamped. If a customer submits a semi-structured or unstructured log file to the index search tool, instant gratification is not possible. These indexed search tools cannot ingest and index semi-structured and unstructured log files without first having to transform these log files into a consumable format. To effect this log file data transformation, a user will first need to invest in upfront programming and REGEX coding to transform data into a structured format that can be ingested by the indexed search tool.

The inherent difference between Glassbeam's machine data analytics solution and these indexed search tools can be expressed in terms of "Indexed Data Search" and "Data Analysis". The indexed search tool approach is based exclusively on ingesting, indexing, and loading structured, time-stamp data into its proprietary data store. Once indexed, an indexed search tool user can submit queries for matching strings against the submitted indexed search query parameter. Each search query requires asynchronous parsing performed in RAM/Cache resulting in noticeable delay in large log files.

Glassbeam enables both indexed search capabilities and data analysis. Glassbeam ingests and transforms semi-structured and unstructured data via SPLTM, its patent pending parsing language. Once SPLTM transforms the data, it is then loaded, in parallel, into up to three different types of data stores:

- LogVault: Data store for the original raw log files.
- Cassandra: Columnar data store for analytics.
- Solr: Indexed data store for search.

The data transformation by SPLTM parsing the data is performed one time; it does not need to be reparsed for indexed queries. Taking a closer look at Glassbeam's three data stores, we'll focus on Cassandra and Solr that enable data analysis and indexed search capabilities. Solr is a widely used open-source application known for its high reliability, scalability and fault tolerance. Glassbeam integrated Solr into its machine data analytics solution to provide the inherent benefits afforded by Solr. Glassbeam's use case for the integration of Solr is not for indexing the submitted log

files. All log files ingested by Glassbeam's analytics engine are parsed to transform the data to an indexed Solr data store and a columnar Cassandra data store. The Solr data store enables Glassbeam users to submit "Full-text" and "Parametric" search queries against the SPLTM parsed data.

The SPLTM transformed data residing in the Cassandra data store enables user data analysis. This transformed data structured as Cassandra Column Families is then available to Glassbeam apps and customer-created apps via Glassbeam's RESTful and API-enabled middleware, named InfoServer. The apps interfacing to InfoServer include both ad hoc and purpose-built "Dashboards" presenting the analytical results of unique use cases applied against the Column Families found in the Cassandra database. These dashboards are both graphical and textual in presentation.

The indexed search tool approach to data analysis is constrained by the limits of its indexed search capabilities. The data stores employed by these indexed search tools are either proprietary (for example, Splunk, SumoLogic, and Loggly) or similar to the data stores used by Solr and Elasticsearch. The data analysis is really based on a set of search queries resulting in frequent, repeated asynchronous parsing when querying indexed data. Submitting search queries against large indexed files can result in noticeable delay in presenting the results.

Another notable difference between Glassbeam's solution approach to log analysis and the indexed search tool approach is found in the capabilities of enabling "alerts" to specific events triggered by a data value or values found in the log file entries. Indexed search tool applications "alerts" are essentially a search query for specific data values representing the occurrence of an event. Just like any other search query, each event query results in asynchronous searching and parsing performed in RAM/Cache, which may result in noticeable delay in large log files.

Glassbeam's approach to "alerts" is based on its integrated "Rules Engine". Glassbeam Rules & Alerts employ Complex Event Processing integrated with a Finite State Machine (FSM), and applied against parsed data. Glassbeam's Rules & Alerts are purpose-built, enabling the definition of specific event rules processed

by the FSM where complex behavior and process workflows need to be modeled and tracked. Glassbeam's finite state machine defines a set of internal states for a process together with the event conditions defining transitions from one state to another. The new state depends on the old state and the input or event prompting a change in state.
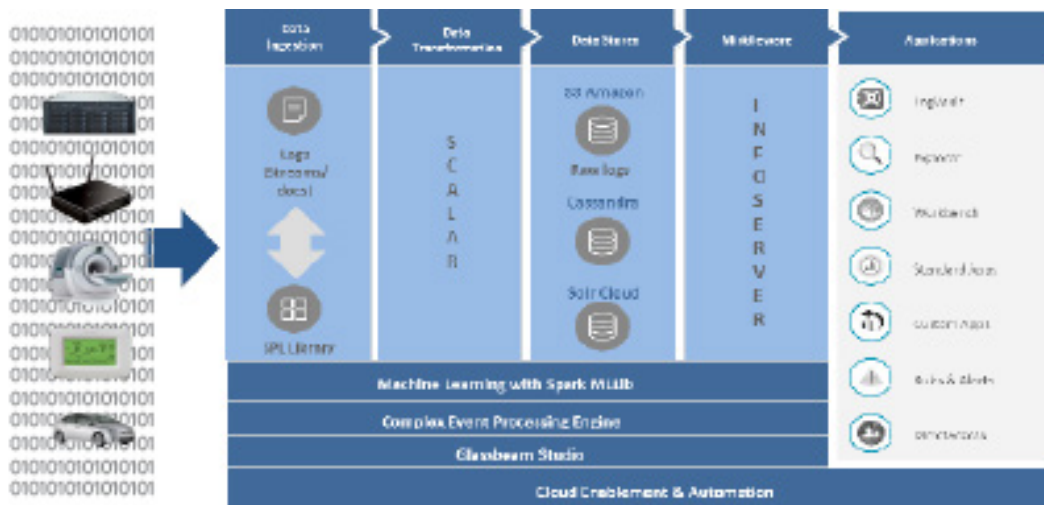
Summary

The subject of price is frequently the initial topic Glassbeam encounters when enterprise organizations compare our solutions approach to machine data analytics to an indexed search tool approach. It's no secret; Glassbeam is more expensive than these indexed search tools when price is measured as "cost per GB". What's not readily apparent in a casual price comparison are the features and functionality Glassbeam and indexed search tool applications include in their cost per GB price.

The data volume pricing for indexed search tool applications includes ingesting, indexing, and loading well-known log file formats with time-stamp data into their proprietary data stores. The not so obvious caveat is that the ingested data must be structured in the form required to index the log files.

Thus, the COST NOT INCLUDED in the indexed search tool application price is the

cost of transforming a log file whose format cannot be directly ingested and indexed by the log analysis application tool. There is a real cost required to transform log files to a consumable format. The sum of these costs includes writing programs and coding the REGEX required to transform the data structure to a compatible format.

Glassbeam's cost per GB is higher than these indexed search tool applications because Glassbeam is offering an analytics solution rather than just an analytics tool. Glassbeam data volume pricing includes ingesting and transforming semi-structured and unstructured data, and then loading the transformed data in up to three different types of data stores.

Indexed search tool application users are on their own to install and use the indexing tool unless they purchase professional consulting services or hire qualified employees. These tools also require the need for someone to write complex queries, create dashboard charts and reports; and maintain them. The creation and maintenance of these queries and dashboards falls beyond the task of installing the indexed search tool application; this is an ongoing function requiring one or more full-time qualified employees.

Glassbeam's solution does not require additional staff for installing and implementing our machine data analytics solution. Glassbeam does offer incremental professional services for our customers preferring a complete end-to-end solution without having to commit additional human resources. Our customers frequently ask for help with the following types of professional services:

- Data Scientists to extract additional custom analysis from data stores via APIs provided by InfoServer.
- Programmers to exercise the InfoServer API to data stores for creating custom in-house apps.

The value of Glassbeam's solution approach is not based on "per GB" pricing, but rather based on TCO. Glassbeam is confident in the value of its pricing to offer prospective customers an opportunity to measure our price/performance in terms

of ROI. We welcome the opportunity to discuss how the benefits of Glassbeam's machine data analytics solution apply to our reader's environment, and to share our approach for quantifying the value of the benefits against TCO resulting in a favorable ROI.

To learn more about Glassbeam's solution approach to machine data analytics and the associated ROI opportunity, please contact Glassbeam at sales@glassbeam or (408)-740-4600.

Contact us at sales@glassbeam.com

Glassbeam, Inc.
5201 Great America Parkway, Suite 360 • Santa Clara, CA 95054
Phone: 408-740-4600 • www.glassbeam.com

Glassbeam, the Glassbeam logo, Glassbeam BI Workbench and Glassbeam Dashboard are trademarks of Glassbeam, Inc. All other trademarks and registered trademarks are the property of their respective owners.