
100 Times Performance Improvement of Big Log Data Analysis



AuriQ Essentia Platform
Whitepaper

AuriQ Systems Inc.
Pasadena CA



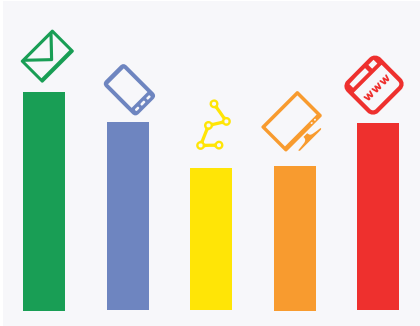
Easing The CMO's Biggest Pain Point

Today's CMO is faced with a rapidly evolving marketing and advertising landscape. There are more channels, more paths, more methods and an ever-growing audience whose behavior is always changing. Through all this, marketers must gain deeper insights into their customers' journey.

Technology advances have made it possible to track almost everything, providing the potential for unprecedented visibility into the customer journey. As a result, marketing has become a data-driven activity in recent years. Improving success, understanding customers, validating spend—each relies on and is augmented by 'big data.'

And yet, this presents CMOs with another challenge: How to cost-effectively manage and make sense of the vast amount of customer data? The sheer volume of big data generated today is often overwhelming. In fact, "data explosion" was cited as the CMO's biggest pain point in a mid-2013 survey, with 71 percent naming it their top concern. In addition, the high cost and lengthy processing times of most current big data analysis solutions often preclude sufficient evaluation of that data.

That's where AuriQ Essentia comes into play. AuriQ Essentia is a brand-new, specialized data processing engine designed for fast, cost-effective analysis of big log data. The AuriQ Essentia platform enables analysis of big log data that's more than 100 times faster than the typical Hadoop-based systems.



Technology advances have made it possible to track almost everything, providing the potential for unprecedented visibility into the customer journey.



The result is actionable intelligence delivered expediently—in time to improve active marketing campaigns.

Why Current Options for Data Processing Fall Short

CMOs have several options for processing and analyzing huge volumes of customer data. Apache Hadoop, for one, is a widely employed, general-purpose platform for analyzing big data. However, using Hadoop to process big log data is like carving up a T-bone steak with a Swiss army knife. Eventually the Swiss army knife will work—but not as easily or as precisely as a good steak knife.

Hadoop relies heavily on writing data to disk. As a result, the processing and analysis speeds are slow and resource-intensive.

Super databases such as Amazon Redshift, which provide a combination of in-memory and from-disk processing, are also becoming popular for big data analysis. They are incredibly fast for aggregating information from large data sets. But as with most relational databases, the problem lies in preprocessing the data into an acceptable format before loading. Preprocessing thousands of semi-structured log files from compressed archives can take days, negating much of the performance gains this type of solution offers.

AuriQ Essentia: Designed for Speed and Scale

By contrast, AuriQ Essentia is designed specifically for speed and scalability. By emphasizing the use of in-memory storage instead of writing to disk, the speed of processing, data cleansing, and analysis of big log data can be up to 100 times faster than other methods.

AuriQ Essentia reduces the cost and time needed to process big log data. As a cloud-based service running on Amazon AWS, any marketer can use AuriQ Essentia without heavy up-front investment costs. A variety of templates are available that enable non-technical users to perform advanced analyses, such as attribution modeling, without any programming skill. AuriQ Essentia reduces big log data into a structured data set, enabling marketers to use familiar tools such as Excel, Qlikview and other business intelligence tools for deeper reporting.

Actionable Intelligence, Delivered In Time to Improve Active Campaigns

Log files are automatically created and updated when a computer, system or user (such as a customer) performs an action. Big log data, then, can be a goldmine of important information to help drive key marketing and business decisions.

However, ad server logs often run into billions of rows of data on impressions, clicks and more. Attempting to combine them with web server logs (which are also often in the billions of rows) is challenging, to say the least. Because of the time it takes to properly correlate and analyze this data, it's difficult to extract actionable intelligence.

AuriQ Essentia enables advanced marketers to tap into the goldmine of information that exists within big log data. AuriQ Essentia provides the unique ability to “match” users from one data log to another, thus creating a complete picture of each customer's journey. With AuriQ Essentia, data analysis that once took weeks now takes only days, and what once took days can now be accomplished in hours. The result is actionable intelligence delivered expediently—in time to improve active marketing campaigns.

A Three-Step Process Toward Actionable Intelligence

AuriQ Essentia uses a three-step process to enable fast, cost-effective analysis of big log data for marketers and digital advertisers.

1



Upload the Data

High volume log data is loaded to Amazon S3 or fetched via FTP. Tell AuriQ Essentia what you're looking for in your final data set, such as unique segments and relationships.

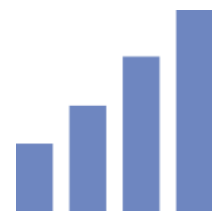
2



Crunch the Numbers

AuriQ Essentia preprocesses raw log data using its exclusive approach, which provides super-fast performance.

3



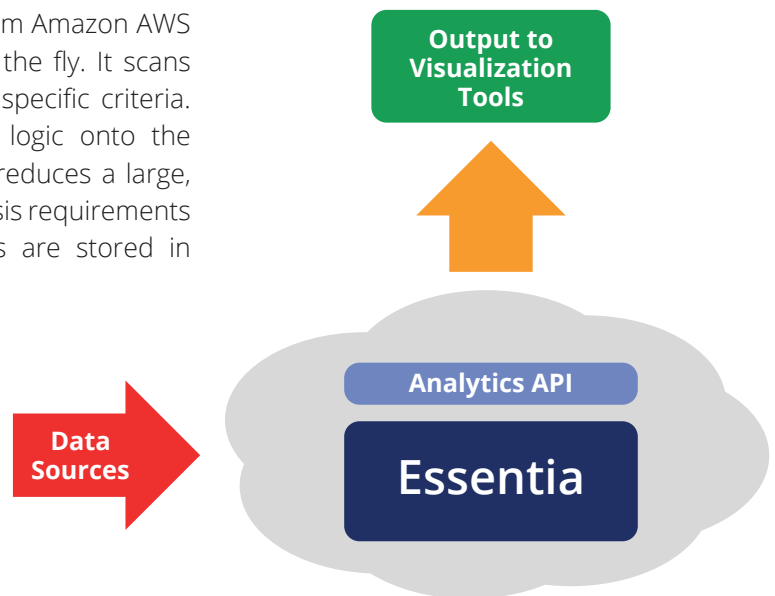
Receive Results

With your specified data set extracted, you can run queries in memory. The results are delivered back to the server for use in your preferred visualization or reporting tool, such as Excel.

How Essentia Works

Single Pipeline Preprocessing

AuriQ Essentia obtains compressed big data log files from Amazon AWS S3 or other storage locations, which are unzipped on the fly. It scans and extracts records from logs that match the user's specific criteria. Preprocessing tools apply filters and transformation logic onto the extracted data. Unlike other solutions, AuriQ Essentia reduces a large, complex data set to a smaller data set that fits your analysis requirements in a continuous, fluid pass. No intermediate outputs are stored in preprocessing; only the target data sets.



Optimized for Logs

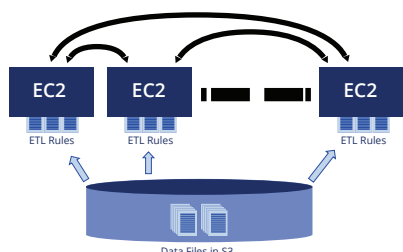
Unlike other solutions, AuriQ Essentia's code is optimized for log processing. The streaming unzip ability lets you work with compressed log files without the need to completely decompress the files before processing can begin. During preprocessing, AuriQ Essentia rapidly removes duplicate and unusable data. Powerful tools match data based on your specifications from different logs. AuriQ Essentia works with any delimited file type.



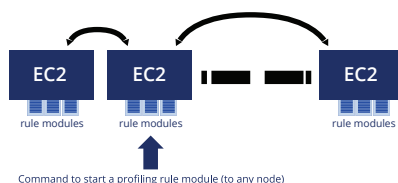
Highly Parallel Processing

All AuriQ Essentia functions are designed to work in parallel across multiple server nodes, such as...

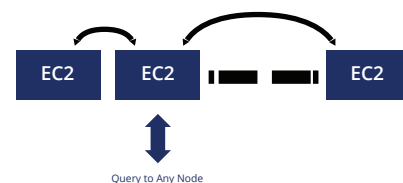
Parallel Preprocessing (ETL)



Parallel User Profiling

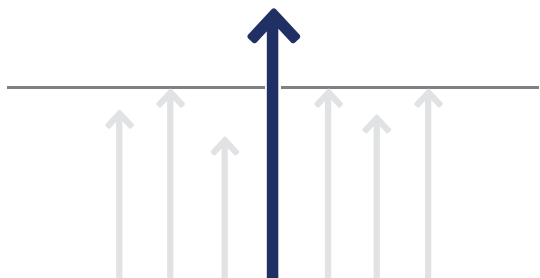


Parallel Query Processing



Fast Query Processing

AuriQ Essentia's optimized code for query processing is extremely efficient. The integrated in-memory database eliminates latency that results from disk I/O. A basic command line interface lets you filter, segment, and aggregate data in memory. You can run thousands of queries against target data sets at blazingly fast speeds. The query processing engine accepts custom modules for complex calculations or scoring for attribution analysis, cluster modeling, and more.



Groundbreaking Performance

AuriQ Essentia achieves significantly faster ETL processing and 100 times faster query processing compared to Hadoop-based systems. Its cloud-based architecture makes it easy and cost-effective to increase performance by adding server nodes.

Flexible Integration with Existing Tools and Processes

AuriQ Essentia can be employed in a variety of use cases depending on your needs:

As a Preprocessor

AuriQ Essentia can be used as an efficient preprocessor for high volume log files. In many cases, the biggest bottleneck to processing log data is the application of ETL (Extract, Transform, Load) in preparation for data upload. Optimized preprocessing tools can cleanse, normalize and load the log data into a traditional relational database management system or a Hadoop Distributed File System.

As a Complementary Tool

The modular architecture of AuriQ Essentia makes it easy to utilize specific tools within the platform, from ETL to aggregation to the application of complex query analysis. Each feature can be used on its own, provided that the input data source is of a delimited format. APIs allow you to import and export data into other applications, and automation tools provide easy scheduling of recurring processes.

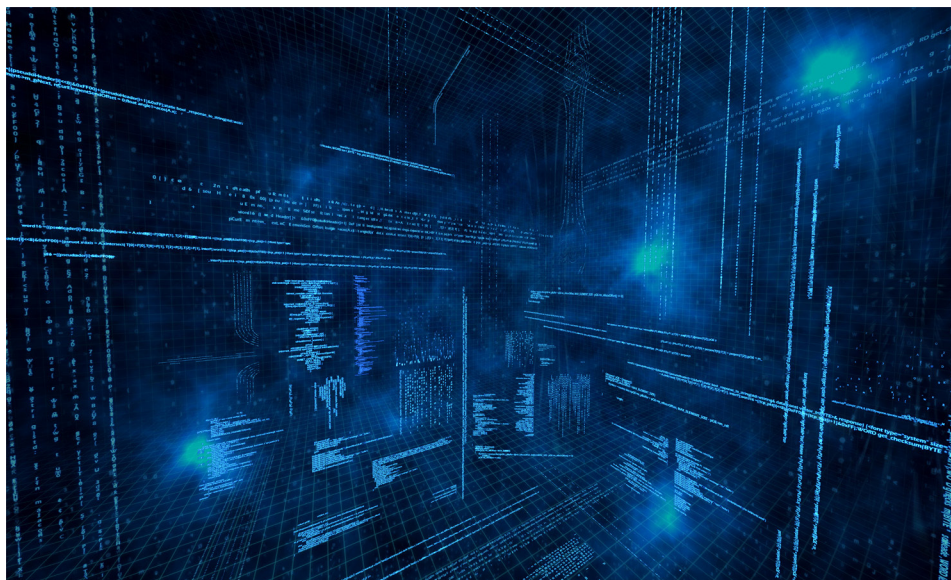
As a Complete Solution

When log files are the primary source of big data for your organization, AuriQ Essentia is an ideal alternative to running a Hadoop cluster. It eliminates the high overhead requirements as well as the need for specialized expertise to operate and maintain. In most cases, customers who have switched to AuriQ Essentia pay one-tenth the monthly costs of operating a typical Hadoop system.

AuriQ Essentia: Benefits at a Glance

In summary, AuriQ Essentia is a revolutionary new technology that helps marketers cost-effectively transition from data overload to actionable marketing intelligence with speed and scalability. With AuriQ Essentia, you can:

- ✓ **Process terabytes—even petabytes—of big log data quicker and at lower cost than other tools.**
- ✓ **Analyze big log data faster than typical Hadoop-based systems.**
- ✓ **Eliminate the need for up-front technology investments and keep ongoing costs low, because AuriQ Essentia is cloud-based.**
- ✓ **Enable non-technical users to perform advanced attribution modeling without any programming skill.**
- ✓ **Liberate IT staff from tedious processing tasks, enabling them to focus on more strategic projects.**
- ✓ **Start using the tools easily. Just upload log data and AuriQ Essentia will clean, join, analyze, and output actionable data based on your specifications.**
- ✓ **Easily import actionable data into familiar reporting tools, such as Excel.**



Want to learn more about AuriQ
Essentia's big log processing
capabilities? Do you have a project
in mind already?

**Please contact us for a free consultation. We'll
assess your data analysis needs and help you
determine if AuriQ Essentia is right for you.**



AuriQ Systems, Inc.

301 North Lake Ave, Suite 1001

Pasadena, California 91101

(626) 564-2781

essentia@auriq.com