CLABBY ANALYTICS

Research Report

Syncsort: Way More Than Meets the Eye!

Executive Summary

The company's name indicates that it is in the "sorting" business – and it suggests that it offers its customers ways to rearrange data set records, merge data sets and copy files. But there's way more to Syncsort than is revealed by the company's name: Syncsort is the maker of Big Data Hadoop, Big Iron to Big Data analytics, mainframe, cloud and Linux/-Windows/Unix solutions that speed processing and simplify management.

A closer look at Syncsort product offerings shows that the company offers:

- Big Iron to Big Data Analytics that includes real-time streaming of mainframe log data to Splunk Enterprise;
- Hadoop-focused Big Data products that include Hadoop extract, transform and load (ETL) services, data warehouse and mainframe to Hadoop offload facilities; Hadoop sort acceleration and Hadoop connectivity facilities. Further, the company's ETL products and Hadoop offerings can run in the Amazon Elastic Compute Cloud (EC2) as well as Amazon Elastic MapReduce (EMR).
- Sort, copy, compress software products and for mainframes; and,
- Sort and ETL products, as well as data migration and system optimization services, for Linux, Windows and Unix environments.

When it comes to the ETL process, we believe that, if it is feasible, data should processed in the same environment that was used to create (and now house) that data. However, there are situations where data extraction, transformation and loading makes sense (for instance, when a data lake needs to be created to provide scale for large, Big Data Hadoop environments). In this situation, it is more convenient to house disparate data at a common source. Syncsort offers products that can accelerate data sorting, ETL and data integration for both mainframe and Linux/Windows/Unix environments.

A closer examination of Syncsort product offerings also shows that Syncsort has developed deep expertise in the processing of Hadoop data. Further, the company now offers operation intelligence facilities that can be used to enhance mainframe security, optimize workload balancing, and minimize/prevent downtime due to system failures.

In this *Research Report*, *Clabby Analytics* takes an in-depth look at Syncsort's market position, its product offerings and the company's recent acquisitions. And what we find is a company that has preserved its "sorting" market niche while successfully expanding into new markets such as the Big Data and cloud marketplaces. With solid sort facilities – and with developmental strength in one of the hottest growth markets in Big Data (Hadoop database processing and management and Splunk Enterprise) – we see Syncsort well positioned for future growth in both the mainframe and distributed server marketplaces.

Company Background

Syncsort was founded in 1968 when seven graduates of New York University and Columbia determined that they could build a better, more efficient sort utility that occupied less disk space than the sort utility offered by IBM. The company's sort utility helped pioneer the market for 3rd party mainframe software (a relatively new concept in that day because most mainframe software at that time was developed by IBM). Syncsort's sort utility now runs on approximately 50% of installed mainframes.

As time passed, Syncsort continued to improve the performance of its sort utility, while introducing other products that could help extract, transform and load data (in effect, by the early 1990s, Syncsort had become an early pioneer in the ETL marketplace).

The next logical step for Syncsort, now that it had facilities that enabled data to be sorted at high speed as well as be extracted, transformed and loaded on other servers, became data collection, processing and integration. In 2011 and 2012, the company saw market demand for ETLing data into Hadoop databases – and adjusted its market strategy to serve the Hadoop/Big Data community. Syncsort's Big Data strategy focuses strongly on building tools to assist with Hadoop data integration and processing. At the same time, Syncsort is collaborating with other software makers (such as Splunk for visualization of mainframe logs in real-time) to broaden and deepen its Big Iron to Big Data product offerings.

Syncsort is also growing through software acquisitions. In 2013, Syncsort acquired Circle Computer Group – the company's first acquisition since its founding. Circle's software enables Syncsort to migrate IBM Information Management System (IMS) data to IBM's DB2 database on the mainframe z/OS operating system. Further, earlier this year, Syncsort acquired UK-based William Data Systems, a maker of advanced network monitoring and security software products for mainframe environments. Syncsort's acquisitions augment its Big Iron to Big Data strategy

> ***After almost 50 years in business, Syncsort now boasts that its software runs on almost half of all mainframes in the world – and its ETL products are also heavily used in Linux, Windows and Unix environments. The company has shown a consistent history of logical innovation, a natural progression moving from basic sort to ETL to Big Data migration and management.***

Organizational Structure

Syncsort is headed by CEO Lonne Jaffe. Josh Rogers (the company's president), owns responsibility for sales, development, marketing and engineering. Reporting to Rodgers are three general managers who run the company's mainframe software, Big Data and sort initiatives.

Competitors

In the Big Data arena, Syncsort's primary competitors are data integration vendors including Talend and Informatica.

Syncsort's Mainframe Business Unit

The focus of Syncsort's mainframe business unit is to help customers find ways to lower mainframe costs and facilitate real-time data from Big Iron to Big Data analytics. This unit offers: 1) Syncsort MFX for high performance sort; 2) Syncsort Ironstream™ for security

and operational intelligence; 3) Syncsort DL/2 and VS/2 for data migration to DB2; and 4) Syncsort Zen for z/OS network management and security

Syncsort MFX

Following in the founder's footsteps, Syncsort MFX product family (MFX for z/OS, ZPSaver Suite™, MFX PROCSort, , MFX PipeSort MFX for z/VSE) continues to deliver fast, optimized mainframe sort, copy, compress and join technology. Syncsort MFX helps mainframe users reduce billable CPU time; support more processing in an LPAR; meet service level agreements; and helps delay the need to purchase more hardware (through more efficient use of existing hardware).

Most notable in the MFX family is the ability of Syncsort MFX to outperform IBM's DFSORT (by 20-40%), and the ability to offload expensive z computing cycles to less expensive zIIP processors.

Also noteworthy, Syncsort MFX PROCSort can be used to provide high performance sort specifically for SAS applications.

Ironstream™

A few years ago, *Clabby Analytics* started writing reports (here, here and here) about a new generation of systems/storage/network management tools that were capable of analyzing massive amounts of enterprise log data and other data streams, making it possible to more efficiently troubleshoot computing problems. Further, we described how these same kind of analytics tools were being used to strengthen enterprise security (here and here). Syncsort's Ironstream™ software (amplified through the acquisition of Williams Data Systems), combined with Splunk software (a software environment that helps enterprise information technology [IT] administrators visualize operational data), enable Syncsort customers to search, monitor, analyze and visualize machine data).

Syncsort's Ironstream™ is used to collect mainframe log records in real-time (such as SMF, SYSLOGs, Log4j and other data streams), and then translate/transform/prepare that data for analytics. This data is then forwarded to Splunk's Enterprise and Cloud environments for real-time analysis of operational data – and is presented to administrators using Splunk visualization tools. This offering is highly scalable – able to handle the analysis of billions of SMF/-SYSLOGS per day with minimal CPU usage. Data that is analyzed can be used to help identify potential security risks and threats and assure application performance – using predictive analytics to recognize potential problems before they occur)

DL/2

Through its Circle Computer Group acquisition, Syncsort was able to improve its data migration offerings to now include the ability to integrate IMS/DB data seamlessly with DB2 data. Syncsort's VS/2 data migration environment enables the rapid migration of VSAM data to DB2.

Syncsort Zen

With the William acquisition, Syncsort added products, technology and expertise related to z/OS network performance and security. Zen offers a selection of specialized components

that provide network performance management, monitoring, alerting, tracing, reporting and security. This software can now be used to power Ironstream™ by providing this same network log data to Splunk for real-time network security and management

The Big Data Business Unit

The focus of the Big Data business unit is to help database administrators move data into data lakes where mainframe and distributed data can be easily shared. Data can be easily ETLed to a data lake by taking advantage of Syncsort's well-established, well-regarded mainframe and distributed ETL tools. But resolving data format issues (such as blending ASCII data with EBCIDIC data) can be a bit more challenging…

For enterprises looking to run Hadoop MapReduce applications, Syncsort offers a wide variety of data movement/data integration tools designed to offload data and expensive workloads from a number of sources (mainframes, Linux, Windows and Unix environments) and make that data available in a consistent form to data scientists.

Syncsort Hadoop solutions include:

- Hadoop Data Integration;
- Data Warehouse Offload to Hadoop;
- Hadoop Sort Acceleration;
- Mainframe Offload to Hadoop;
- Hadoop Connectivity;
- Ironcluster® Hadoop ETL for Amazon EMR; and,
- Ironcluster® ETL Amazon EC2 Edition

The big challenge when ETLing data is that traditional ETL architectures (distributed systems) are facing scalability issues as they deal with massive amounts of Big Data. Syncsort's DMX-h Hadoop ETL Solution helps IT and database administrators deal with this scalability issue by enabling any data source to be connected with any data target; by enabling MapReduce ETL to take place without coding; by enabling the scripting of common ETL tasks; and by optimizing the performance and efficiency of individual nodes.

Syncsort's Data Warehouse Offload to Hadoop does exactly what the product name describes: it enables IT administrators to offload data and ETL workloads from data warehouses into Hadoop databases. Using this offering, IT and database administrators can significantly reduce batch processing windows, keep more data readily available, while freeing-up data warehouse capacity.

Hadoop Sort Acceleration makes it easier to sort data for MapReduce applications. Using the DMX-h Hadoop Sort Acceleration Solution, database administrators can accelerate Map-sort and Reduce-merge operations, and optimize the performance and efficiency of individual nodes.

Mainframe Offload to Hadoop makes use of Syncsort DMX-h ETL to simplify data ingestion, translation, processing and transfer – making it easy to move mainframe data into Hadoop processing environments. Mainframe data can be ingested directly from the

mainframe into a Hadoop database; that data can be translated on the fly, and ETL jobs can be created visually without the need for coding. Mainframe JCL batch workloads can also be moved from the mainframe to Hadoop. DMX-h Hadoop ETL Edition augments mainframe connectivity to Hadoop databases. For enterprises interested in running ETL in the cloud, with or without Hadoop, Syncsort offers two products: 1) Ironcluster Hadoop ETL for Amazon EMR; and, 2) Ironcluster ETL for Amazon EC2.

Summary Observations

The company name – Syncsort – is a misnomer. Syncsort was once all about sorting – but now it is so much more. Syncsort products sort data, move data, migrate/convert data, and format data to serve MapReduce Hadoop applications in integrated large data pools. The company has almost six decades of experience in mainframe environments, and four decades of experience in converging mainframes with distributed systems. It knows how to move, convert and manage data from all platforms both in batch and in real-time mode to support large scale Big Data Hadoop deployments.

In the past, Syncsort's approach to product development had been largely internally focused. But in 2013, Syncsort made its first acquisition, Circle Computer Group (to add additional data conversion facilities to its product line) – and followed that acquisition by acquiring William Data Systems (for advanced network monitoring and security software products for mainframe environments). This is a welcome change from the Syncsort of old – these acquisitions enhance other Syncsort product offerings while at the same time allowing the company to expand into new areas (such as operational intelligence). We'd like to see Syncsort acquire even more software companies and expand its portfolio accordingly.

Another example of how Syncsort has moved from its not-invented here mentality to a more market-responsive, adaptive company is evidenced by its relationship with Splunk. By forming a tight relationship with Splunk, Syncsort customers can now get access to a rich capture, convert, push environment made up of Syncsort Ironstream™ technology for real-time data capture that is complemented by Splunk's world class visualization and analytics environment. Syncsort also has close relationships with several professional services providers – enabling the company to simplify deployment and integration for its customers.

There have been some big changes in company developmental philosophy and in company strategic direction over the past two years at Syncsort. For enterprises looking for an advanced data movement, data integration environment that can handle large Big Data Hadoop database requirements and Big Iron to Big Data challenges, Syncsort is worth much closer scrutiny.

Clabby Analytics
http://www.clabbyanalytics.com
Telephone: 001 (207) 846-6662