

# Global IDs makes 'big data' MDM play via the cloud and Hadoop, touts profitable growth

**Analyst:** Krishna Roy

11 Feb, 2013

Global IDs remains on a mission to focus on large-scale information integration. The 12-year-old company is looking to address this issue, along with those associated with the management of semi-structured data in large enterprises, in its next major release through the use of the Hadoop open source data-processing framework and Amazon's cloud. Global IDs' initial 'big data' master data-quality and profiling push within its forthcoming Master Data Governance Suite version 8 isn't likely to be its last. The company is mulling the delivery of a so-called Business Ecosystem Optimization offering that is designed to provide visibility into other types of data and databases, including NoSQL engines. Having reportedly grown revenue by 100% in 2012, when it racked up a sixth year of profitability, Global IDs plans to increase sales by at least the same percentage in 2013. It is also looking to bulk up in a bid to better address direct sales opportunities.

## The 451 Take

Large enterprises in financial services and telecommunications – two verticals in which Global IDs has a longtime and growing presence – tend to be the early adopters of Hadoop, NoSQL databases, streaming data and other big data environments. It therefore makes sense to provide data-governance capabilities to govern these new emerging environments. However, Global IDs is not the only player looking to serve these new opportunities (see below), which are likely to increase as more big data environments go from pilot deployments and project-specific rollouts into enterprise-wide deployment. Nonetheless, the company seems primed for further profitable growth in 2013, which suggests it is thriving as a smaller vendor

among data management giants. But market consolidation continues apace, and we still can not rule out a takeover, given the company has had acquisition offers before.

## **Context**

Now in its seventh year of operation as a software provider, having started out in 2001 as a data management and integration consultancy, Global IDs is preparing to deliver the next major release of its Master Data Governance Suite sometime in March. Built on an agent-based platform written in Java that uses parallel processing in a distributed environment, one end game behind the next release is to serve customer requests to roll out its data management software more broadly across the enterprise. The company decided to draw on the elasticity of Amazon's EC2 cloud environment to meet the need for wider enterprise deployments. It may also serve up deployments on other clouds, such as AT&T's Synaptic Storage as a Service and Apache's OpenStack open source hosting environment, in the future.

The forthcoming release was also crafted in response to customer requests to bring NoSQL environments, such as the Apache Cassandra and MongoDB databases, into the master data management (MDM) mix and, furthermore, manage and process streaming data using Hadoop. Additionally, Data Governance Suite version 8 contains a master data portfolio component designed to enable users to examine all master data (customers, products, stores, etc.) by profiling the data as a portfolio of assets using Hadoop's processing capabilities. Version 8 is embedded in Cloudera's Hadoop, and can talk to the Hadoop File System because its agents speak Pig Latin - the textual language used to program data transformation in Hadoop. As a result, it is designed to enable large-scale data quality so that users can understand the quality of master data at a very granular level.

A potential next development would be to expand visibility into data to the entire ecosystem by supporting new environments such as NoSQL databases, as well as providing support for a wider variety of data domains within these environments. Currently a speculative initiative dubbed Business Ecosystem Optimization, the end game is to enable enterprises to more effectively optimize enterprise data by providing them with analysis and insights into the nature of the data and its location.

On the business front, courting Fortune 50 companies, which need to automate data management and gain cost benefits by eliminating the need for the manual management of hundreds of

thousands of databases, remains the order of the day. Global IDs' software stack uses specialized agents for tasks such as cleansing, profiling and mastering data. It takes a semantic-based approach to data management that also involves the scanning of database environments to gain a detailed profile of the data within them, including the metadata, as well as data mapping at the schema and domain level.

The company now claims to have less than 25 paying customers, compared with approximately 15 paying accounts in March 2012. It reportedly continues to land enterprise deals in the seven figure-range, compared with six-figure deals a couple of years ago. From a revenue perspective, direct sales account for the lion's share of business, bringing in 60% of sales, with the remainder coming from the channel. However, partners are reportedly generating more business from a volume perspective, accounting for 60% of the deals the company lands. That said, the direct/indirect deal mix may change in the coming years as the company builds out its direct sales operations. The game plan is to take headcount to more than 100 employees, from 67 staff currently, with the bulk of the recruitment drive focused on sales.

## **Competition**

Global IDs' Master Data Governance Suite is employed for a variety of use cases, including mastering, profiling and cleaning master data, integrating data, and metadata management. The company contends that the suite's agent-based approach enables it to serve these tasks in a single, fully integrated stack, which is hard to build using a homegrown approach and difficult to buy from another data management purveyor because rivals have adopted a 'buy versus build' strategy, resulting in a collection of tools under one brand, and often a consulting component, as well.

We have certainly noted a heightened level of M&A activity in the data management sector. This was most recently exemplified by Actian's successful bid to buy database and data-integration specialist Pervasive Software. IBM and Informatica – data management providers with a focus on large enterprises and, therefore, competitors to Global IDs – have also played a part in consolidating the sector. Informatica's latest M&A play involved reaching for UK data-quality player Data Scout Solutions Group. The strategic tuck-in acquisition followed in the wake of the purchase of product-oriented MDM firm Heiler Software in October 2012. Big Blue also continues to ink acquisitions to augment its existing data management portfolio, as well as shore up its ability to handle semi-structured data from a management and analysis perspective.

We would also point out that IBM has a big data integration and governance strategy that is supported by a broad portfolio of software and services. Informatica has focused on better

supporting big data for data management scenarios such as data governance for a couple of years now – as evidenced by the last major Informatica 9.5 integration platform release.

Additionally, we think SAS Institute's focus on the Fortune 50, together with its ability to deliver data quality and integration in support of data governance associated with analytics, makes it competitive to one degree or another. That said, Global IDs' forthcoming release does deliver differentiation in support of emerging data sources, which we believe helps its case as a smaller pure play among the data management leviathans.

Global IDs is also one of several other players looking to serve the needs of emerging environments associated with big data. Although not as focused on data governance and MDM as Global IDs, BI vendor Pentaho recently introduced a user-friendly data-integration, modeling and analysis application for Hadoop and NoSQL databases. Talend – another commercial open source firm – is also operating in the big data management fray on a number of fronts, including data quality, profiling and integration scenarios associated with Hadoop.

## **SWOT Analysis**

### **Strengths**

Global IDs appears to be holding its own as an independent player in data management. It is exhibiting stellar growth while remaining profitable, as well as forging into new arenas much needed by large enterprises, which are now starting to address the data-governance issues arising from big data.

### **Opportunities**

Existing customers are low-hanging fruit for Global IDs' upcoming release, which should also open the door to new deals where data governance is required at scale as well as new emerging data sources arising from semi-structured information and streaming data.

### **Weaknesses**

Although the company's software stack does bring undoubted cost benefits by automating many data management tasks that would otherwise have to be carried out manually by IT, it doesn't come cheap. A cloud delivery model isn't for everyone.

### **Threats**

IBM and Informatica are potent rivals that may not offer the same capabilities in quite the same way, but are incumbent providers in large enterprises. With data management sector consolidation ongoing, we wouldn't rule out an acquisition of Global IDs, which has had offers before and turned them down.

Reproduced by permission of The 451 Group; © 2013. This report was originally published within 451 Research's Market Insight Service. For additional information on 451 Research or to apply for trial access, go to: [www.451research.com](http://www.451research.com)