

# BMC Software's batch-job juggernaut gets hip with Hadoop support

**Analyst:** Michael Coté

7 Nov, 2013

The domain of managing scheduled execution of arbitrary workloads – batch-job management – can seem like a painfully boring field. What's curious, though, is that a major component of the ever-popular Hadoop is exactly that: kicking off a set of processes distributed over multiple machines, and orchestrating management of the resulting data set.

As Hadoop use continues to skyrocket, large companies are looking for better ways to manage those Hadoop 'job runs.' These companies have the usual enterprise requirements for control, governance and manageability that off-the-shelf Hadoop may not accommodate. It's little wonder, then, that BMC Software has added support for Hadoop to the batch-job framework of its Control-M management platform – largely driven by an upswing in customer interest.

## The 451 Take

Interest in Hadoop is quickly converting to actual mainstream use: BMC says that of its 3,000+ Control-M customers globally, between 60% and 70% are interested in, if not looking at proofs of concept for, Hadoop. Another proxy for Hadoop growth is our own revenue estimate for Cloudera, the provider of Hadoop-based software, support and services, at just north of \$120m for 2013. As Hadoop enters the mainstream, classic enterprise control and governance levers will be in high demand, especially from larger companies. BMC has been a top, if not the leading, vendor in this space for ages, and has an extensive customer base that will want classic enterprise management over Hadoop. Providing Control-M customers with the ability to deploy Hadoop as well as managed internal services should give BMC one more way to retain its market position.

## Context

Houston-based BMC has been in the IT management business for decades, and is often referred to as one of the Big 4, the other three being HP, IBM Tivoli and CA Technologies. BMC is a pure-play systems management software company that aims to be a one-stop shop for everything large enterprises need to manage their IT, on both sides of the firewall. BMC services both mainframe and 'distributed' (mainframe-speak for x86 platforms, Windows, Linux and Unix) computing – with a rough split of 60% distributed and 40% mainframe, based on top-line revenue reported for the quarter ending June 30, 2013. The company acquired Control-M in 1999 when it purchased New Dimension Software, and has maintained a steady foothold in the batch-job management market.

Rumors are that Control-M experienced a surprising growth spurt in recent years. The company more or less confirmed this, saying its use has been growing at twice the average market rate. BMC products are often thought of as mainframe-heavy, the knock being that tooling a potentially shrinking ecosystem (which has managed to remain arguably steady despite prognostications) is bad business, long-term. The company said that Control-M is used with a mix of mainframe and distributed compute 60% of the time, only used with distributed 39% of the time, and used exclusively for mainframes 1% of the time.

BMC went private as of September 10, 2013 in a \$6.9bn deal with Bain Capital, Golden Gate Capital, GIC and Insight Venture Partners. The market has been waiting to see what BMC will do now that the pressure of quarterly reporting has been removed. If Control-M's claimed growth rate of double the market average is accurate, it would be one of BMC's more attractive cash cows.

## Products

Hadoop is the open source MapReduce framework that has seen steady growth in recent years, fueled by companies' need to analyze large sets of structured and unstructured data. At its core, Hadoop is a batch environment, and so fits naturally into Control-M's bailiwick.

Control-M provides a centralized console for managing and scheduling job runs across ETL workloads and custom-written processes inside companies – for example, syncing customer orders with inventory management systems. Because the execution of these jobs is often core to the day-to-day running of the business, IT departments rely on tools like Control-M to ensure the jobs are run, to report when jobs fail or are taking too long, and to provide all the audit logs required for governance. Larger enterprises often have all manner of systems, new and old, so Control-M

supports mainframe, Tandem, System i, x86, clouds based on VMware and Amazon EC2, and a host of applications such as SAP, Informatica, file transfers and various databases. The addition of Hadoop gives Control-M an even larger pool of platform support.

Control-M's Hadoop support allows users to create and schedule Hadoop jobs, typically custom-written code. These jobs are then distributed across the Hadoop cluster through the phases of data processing and collation known as 'map reduce.' As with all managed jobs, Control-M gives IT admins the ability to centralize, log and otherwise control these Hadoop job runs.

## **Customers and use cases**

BMC says it has more than 3,000 customers for Control-M. These are typically larger enterprises that need to run nightly or weekly batch jobs; for example, to recompute an insurance company's book to determine how much insurance it can sell the next day. Retailers also use batch-job management to run reports on inventory and sales. BMC has noticed many companies using Hadoop for ETL-like runs, collecting data, processing it and then moving it to reports or other databases. Other Hadoop analytic jobs are used to better construct call trees and routing over time for customer-service departments. BMC, of course, sees classic big-data analytics for customers, such as sentiment analysis over reviews and social media.

MetaScale (a big-data spinoff of Sears Holdings) used Control-M in concert with Hadoop to move Sears' batch processing off mainframes, and reduce the overall time taken for those jobs. Another unnamed insurance customer figured out how to add just one field to claims to better understand if the claim could be paid. One of its ETL and batch-job setups took 14 hours to run, but with Hadoop, it took just seven minutes. Other customers, BMC says, are looking at ways to use Hadoop to move workloads off mainframes (to save money), noting that COBOL programmers can be quickly trained on Pig (the Hadoop query language), and get mainframe-equivalent functionality for 'pennies on the dollar' with Hadoop.

## **Competition**

For Hadoop orchestration, BMC says it often sees admins using the open source project Oozie. Cisco Tidal Enterprise Schedule supports Hadoop job management and more general job scheduling. In the broader space of job management, IBM with Tivoli Workload Scheduler and Platform LSF for HPC scenarios, and CA Technologies with Workload Automation AE (née AutoSys), have competitive products and suites.

There are also smaller companies like UC4 and its ONE Automation software suite. As part of a

strategy to expand beyond batch jobs, it recently renamed itself Automic, and declared its intention to focus on automation as well. Often, operating systems and other platforms have job and task scheduling built in: Microsoft Windows, for example, and of course, cron and Unix. These bundled schedulers are by no means as sophisticated as dedicated job management suites, but suffice for 'lower end' uses.

While not direct competitors, companies that focus on big data, such as Pivotal, solve the same problems. There are also efforts from Hadoop distro companies like Cloudera, Hortonworks, MapR and others. After acquiring Infochimps in April of 2013, CSC may look to enter the Hadoop job management space - it certainly has the large-enterprise and government customer base that BMC is targeting here.

## **SWOT Analysis**

### **Strengths**

As one of the leading batch-job management vendors, BMC has a broad array of 3,000+ paying customers, many of which appear to be interested in managing Hadoop. Control-M is a mature platform for job management.

### **Opportunities**

There is undeniable interest in Hadoop, which, by anecdotal reports, seems to be spreading into the BMC sweet spot of mainstream accounts. The largest opportunity is in upselling BMC's existing 3,000 accounts, and using Hadoop support as a reason to renew, rather than consider alternative batch-job management suites.

### **Weaknesses**

BMC is not well known for its involvement in Hadoop, and may not be top of mind. Indeed, BMC's reputation as an enterprise shop may seem 'too enterprisey' for much of the Hadoop market, outside of those 3,000 existing Control-M accounts. This may prevent potential new customers from considering Control-M based on fears (valid or not) around cost and the time to get up and running.

### **Threats**

Many parties in the broader Hadoop market are eager to climb the enterprise value stack, no doubt by providing many of the 'enterprise grade' features BMC is shipping. Large vendors with equally large enterprise rolodexes, such as CSC, are highly motivated to enter this market as well.

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)