



Research Report

CA Technologies' Big Data Infrastructure Management – Unified Management and Visibility of Big Data

Executive Summary

CA Technologies recently exhibited new technology innovations, marking its entry into the Big Data marketplace with CA Big Data Infrastructure Management (CA BDIM). CA BDIM is designed to help organizations *manage multi-vendor Hadoop environments across Big Data infrastructures from a single unified view*. Using this streamlined interface, administrators can get detailed insight from data, as well as related systems information such as performance metrics from jobs, nodes and clusters.

CA BDIM unifies data- collection and management across Hadoop environments. Metrics can be gathered from on-premise/off-premise public and private cloud environments. Big Data and disparate Hadoop management domains can all be managed as one common domain through a unified interface.

In this *Research Report*, *Clabby Analytics* takes a closer look at CA Technologies' new entry, CA BDIM. We'll also look at CA BDIM's integration with the company's other products and the use cases that those solutions enable.

Market Background

According to a new market report published in November 2014 by *Transparency Market Research* called "Hadoop Market - Global Industry Analysis, Size, Share, Growth, Trends, and Forecast, 2012- 2018," the global Hadoop market was worth USD 1.5 billion in 2012 – and is expected to reach USD 20.9 billion in 2018, a CAGR of 54.7% from 2012 to 2018. This growth can be attributed to the many benefits that processing Hadoop data can provide (such as faster insights derived from the analysis of very large data sets).

Hadoop is primarily used to easily and cost-effectively piece together multiple low-cost commodity x86 servers into scalable clusters that can readily process large amounts of data in parallel. Hadoop's distributed file system is capable of collecting and storing data from multiple sources in whatever form it has originated – including both structured and unstructured – and resources need not be assigned to configuring data to “fit” into a relational database.

Many organizations use Hadoop as the foundation for a “data lake” where multiple data types – structured and unstructured – are collected from a range of sources, placed in a common repository in their native format to be queried and analyzed at a later point. Hadoop provides the ability to inexpensively store and analyze large data sets and break down organizational siloes, so Big Data initiatives can be launched by enterprises both large and small.

CA Technologies' Big Data Infrastructure Management – Unified Management and Visibility of Big Data

The Competitive Scenario

With the growth in Hadoop, an “ecosystem” has evolved with vendors adding functionality around Hadoop. Some of these vendors include Cloudera and Hortonworks, vendors that provide Hadoop-based enterprise-ready data management platforms; Cassandra and MongoDB, makers of Big Data NoSQL databases; and MapR (a maker of Apache Hadoop software). Further, SAS provides business analytics and business intelligence software for Hadoop environments; and Tableau provides analytics, visualization and interactive dashboards. Many enterprise customers have integrated several of these tools in order to build an integrated, comprehensive data management environment.

From a Big Data management perspective, EMC recently announced a turnkey Hadoop management environment known as the Federation Business Data Lake. This solution takes uses several products provided by the newly formed EMC Federation of Companies (EMC Information Infrastructure, Pivotal and VMware, VCE, RSA). By incorporating VMware Vblocks, EMC Isilon, the Pivotal Big Data Suite and a comprehensive set of services into an integrated common management environment, customers can take advantage of EMC’s end-to-end solution to collect, store and analyze Big Data.

CA Technologies takes a broader approach to the management of Hadoop environments. Instead of building a data collection/management solution around a closed set of vendor solutions, CA Technologies focuses on consolidating multi-vendor Hadoop domains across the entire Hadoop ecosystem – enabling management into a single easy-to-use, unified management and visualization layer. This enables customers to protect existing investments in Hadoop solutions while still providing a common platform for management.

CA Technologies and Big Data: The Strategic Focus

CA defines its Big Data mission:

“Our overall mission is to make it easier for our customers to more effectively manage their infrastructures and to successfully perform analytics that unlock the value of structured and unstructured data”.

The company’s strategy focuses on:

1. Enabling the use of structured and unstructured data for analytics on any platform;
2. The management and simplification of complex analytics environments;
3. Securing and enabling compliant access control to data in analytics databases; and,
4. Delivering analytics solutions that remove the need for special skills and knowledge of the tools and infrastructure

CA Technologies and Big Data: The Architecture

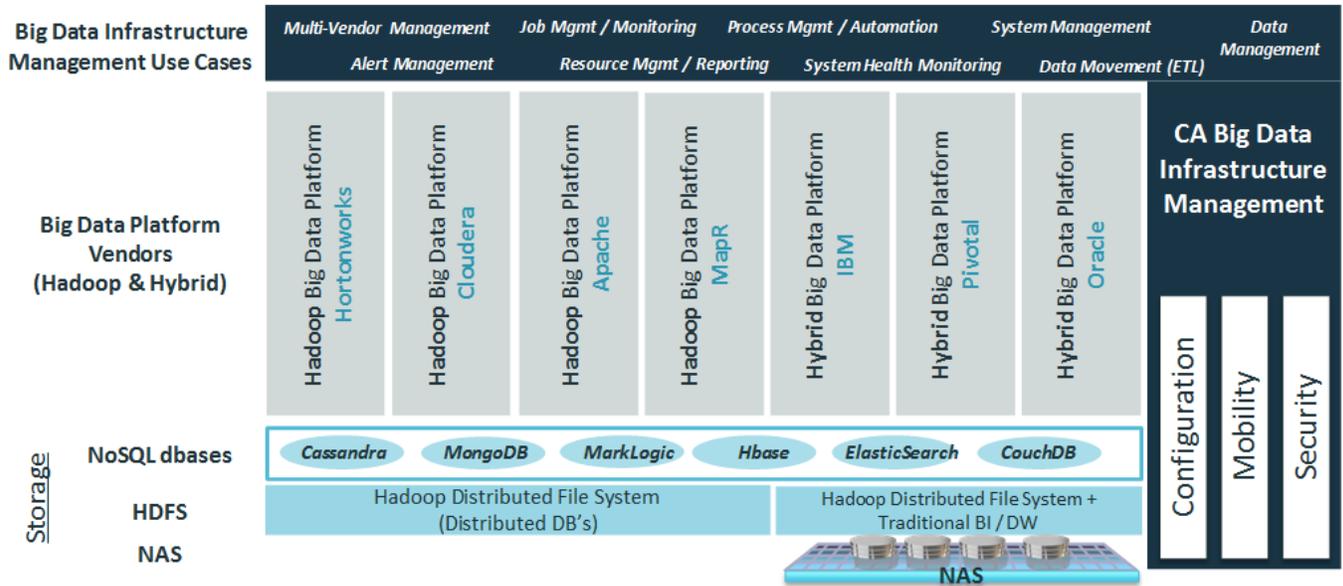
CA BDIM combines a simple, common management interface across multiple management solutions/ Hadoop tools that span enterprise infrastructures.

The architecture (See ***Figure 1***, next page) is designed around management use cases (such as multi-vendor management, job management/monitoring, process management/automation – and the like). Administrators can launch these management use

**CA Technologies' Big Data Infrastructure Management –
Unified Management and Visibility of Big Data**

cases in order to manage Hadoop data environments across various Big Data platforms – as well as to manage associated system resources.

Figure 1 -CA Big Data Architecture



Source: CA Technologies 2015

CA BDIM – A Closer Look

The first release of CA BDIM (available now) supports on-premise Hortonworks and Cloudera. Support for hybrid cloud, as well as Cassandra, MongoDB, Amazon and Yahoo, will be added in the near future leveraging agile development release methodology.

The company expects early adopters to be existing CA Technologies' large enterprise customers in finance, retail and entertainment.

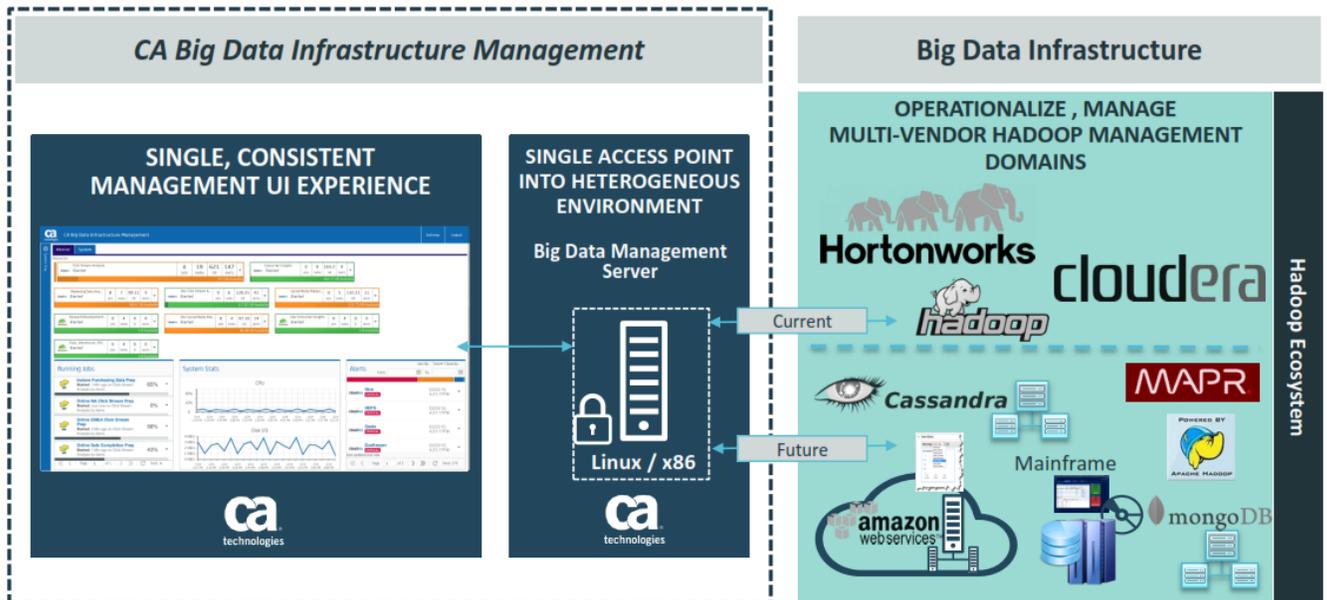
Major features/benefits:

- Monitoring and management of heterogeneous Big Data environments;
- Single unified view of Big Data system performance, metrics and data;
- Automated management and administration of Big Data environments;
- Increased efficiency of labor through improved Big Data infrastructure administration;
- Improved process automation;
- Early detection of system capacity and job bottlenecks to minimize downtime
- Reduced costs of rapidly growing storage environments through use of commodity hardware; and,
- Improved Big Data resource management and reporting.

Figure 2, next page, illustrates CA BDIM working with other vendors Hadoop implementations.

*CA Technologies' Big Data Infrastructure Management –
Unified Management and Visibility of Big Data*

Figure 2- CA BDIM Hadoop Integration



Source: CA Technologies 2015

CA Big Data Management Solutions

CA BDIM is part of a portfolio of solutions aimed at Big Data Management and includes CA Unified Infrastructure Manager (CA UIM) and vStorm Connect Data Streaming for Hadoop (vStorm) enabling several customer use cases.

- CA UIM ‘s central management and a single unified view across heterogeneous distributed and mainframe infrastructure (including Hadoop and Cassandra instances), highlights potential problems that can be diagnosed with Nimsoft probes and “drill-down” capabilities. CA UIM can automatically generate trouble tickets providing end-to-end Big Data management.
- vStorm integrates IBM System z data with data residing on distributed Hadoop infrastructure, providing the ability for businesses to leverage and analyze all enterprise data. Data is transferred from the mainframe and loaded directly into Hadoop (bypassing ETL) on-premise or in the cloud, increasing the agility of data movement from on-premise to off-premise and breaking down siloes between distributed and mainframe infrastructure.

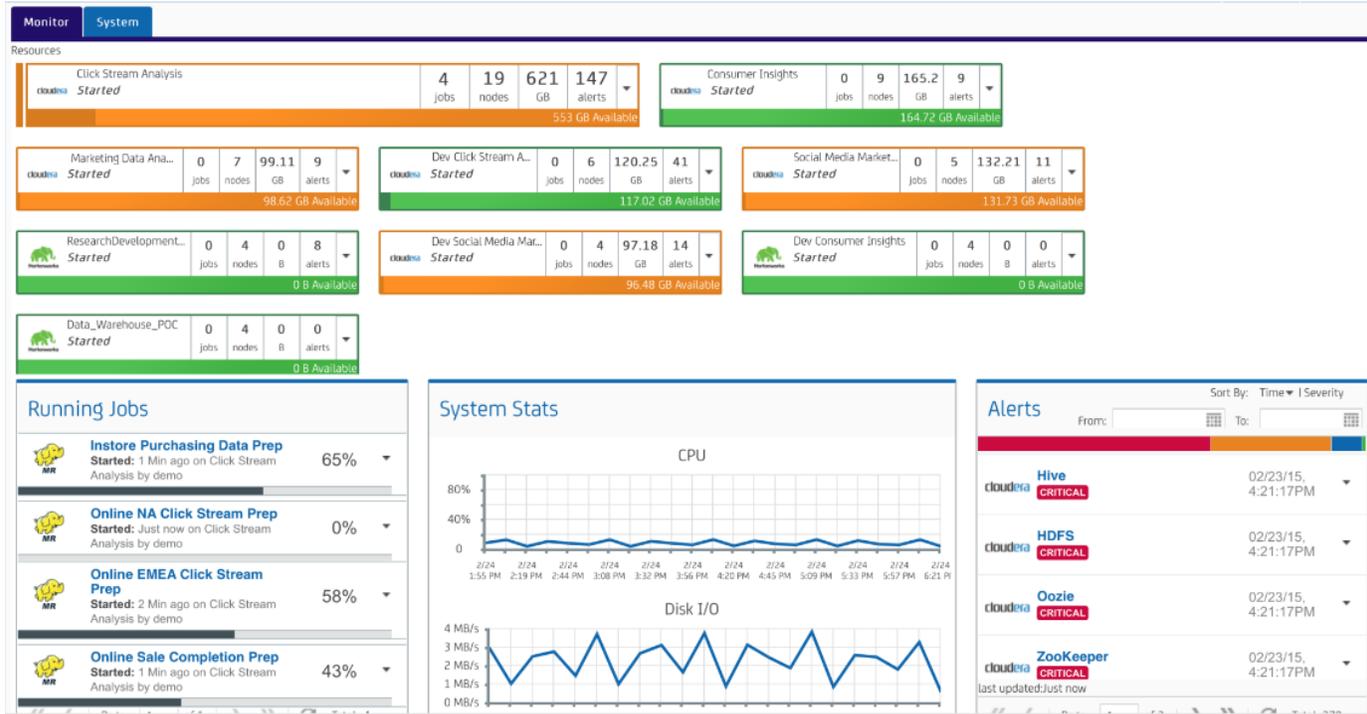
Visualization Layer

CA BDIM’s visualization layer enables central management of multi-vendor Hadoop environments providing a more granular level of insight for data and metric details including jobs, nodes, clusters, system statistics including CPU usage and disk I/O, and alerts from a single unified view. With the added complexity caused by departmental implementations of a variety of Hadoop-based tools and open-source solutions deployed on a range of hardware platforms, central IT organizations have had to step in to provide unified management, analytics and a visualization platform.

CA Technologies' Big Data Infrastructure Management – Unified Management and Visibility of Big Data

With an easy-to-use dashboard interface and analytics that combine data from across the enterprise, CA BDIM can provide more detailed insights that can be derived more quickly. The interface is designed to be intuitive and engaging, with a short learning curve. See [Figure 3](#) below.

Figure 3 – CA BDIM Visualization Layer



Source: CA Technologies 2015

Summary Observations

One thing that differentiates CA Technologies' Hadoop management products from other vendors is that the company offers strong management products across both mainframe and distributed environments (only IBM and a few smaller vendors offer multi-platform, integrated Hadoop management environments). This multi-platform integration element becomes very important as organizations start to gather, store and analyze Big Data that runs in both distributed as well as mainframe environments. Sure, there are a lot of new sources of Big Data including social media, mobile applications and data collected from the cloud – but mainframes still retain a lot of important enterprise data that can and should be analyzed in conjunction with data from these other sources.

CA BDIM can provide an integrated view of Hadoop environments across the enterprise. In addition, CA's Big Data Management Portfolio provides additional functionality that enables a range of management use cases including availability, and performance and SLA monitoring for Big Data, management and optimization of Big Data environments, process/workflow automation and data movement between mainframe and distributed systems.

***CA Technologies' Big Data Infrastructure Management
–Unified Management and Visibility of Big Data***

In addition, CA BDIM allows customers to retain Hadoop tools that they have already acquired and standardized on – while providing a security overlay and visualization layer that displays an integrated enterprise view of system and data metrics. By providing a “manager of managers”, this functionality helps eliminate management siloes and provides a system view as well as a data view. These features enable IT and business users to be more efficient and proactive, identifying performance problems and capacity issues before they impact users.

Buyers interested in consolidating Hadoop environments across enterprise infrastructure for secure, efficient, proactive and cost-effective management of Big Data should closely examine CA BDIM.

Clabby Analytics
<http://www.clabbyanalytics.com>
Telephone: 001 (207) 239-1177

© 2015 Clabby Analytics
All rights reserved
April 2015

Clabby Analytics is an independent technology research and analysis organization. Unlike many other research firms, we advocate certain positions – and encourage our readers to find counter opinions – then balance both points-of-view in order to decide on a course of action. Other research and analysis conducted by Clabby Analytics can be found at: www.ClabbyAnalytics.com.