DATA SHEET

CA Workload Automation Advanced Integration for Hadoop

Apache Hadoop has become the dominant platform as more and more companies start embracing big data and the analytics that come with it to make informed decisions to serve their customers better. Hadoop provides a rudimentary scheduler where users can manage Hadoop jobs to achieve limited automation. A typical customer workload environment comprises a variety of jobs, including Hadoop jobs. There is a need for organizations to closely integrate Hadoop jobs into their workflows along with their traditional jobs to monitor SLAs in real time.

CA Workload Automation Advanced Integration for Hadoop provides CA Workload Automation customers the ability to add Hadoop jobs to their existing workflows along with other traditional jobs and be able to monitor their end-to-end workflows from a single console without having to go through multiple schedulers.

Key Benefits/Results

- **Reduce Hadoop job complexity** and execute in sync with workflows running in the rest of the enterprise.
- **Gain visibility** and take control of Hadoop applications with a single view of workflows across the entire IT environment.
- **Reduce the cost and complexity** of managing mission-critical Hadoop and traditional jobs.
- **Get faster implementation** and more accurate big data analytics from one location.

Key Features

- **Define** Hadoop jobs, like HDFS, PIG, HIVE, SQOOP, Oozie and add to workflows.
- **Comprehensive auditing and flexible reporting** capabilities easily meet corporate governance and regulatory compliance.
- **Administer and monitor** statuses of Hadoop jobs.
- **Native integration** for Hadoop and enterprise applications facilitate an easy way for staff to create, schedule and manage big data business services.

Business Challenges

Limited ability to visualize, monitor and run workflows. A typical end-to-end business workflow may have a combination of Hadoop jobs and traditional jobs. However, Hadoop users currently have to run their Hadoop jobs in a workflow separately from the traditional jobs, limiting their ability to manage or visualize all jobs from one location.

Reduced responsiveness in managing time-dependent, parallel jobs. Hadoop users have limited ability to auto-trigger traditional jobs which may be dependent on a specific job in the Hadoop workflow. They currently have to toggle between Hadoop and CA Workload Automation engines to manage traditional jobs that may be part of the overall workflow. This decreases the ability to manage parallel jobs and increases the time and cost in completing workflows.

Additional effort associated with training users on multiple scheduling engines. The look and feel of Hadoop Scheduler is different from that in CA Workload Automation, and since workflows may require jobs to be scheduled and run from both products, companies may end up with different teams, increasing costs around training and causing an overall reduced responsiveness to the business.

Solution Overview

CA Workload Automation Advanced Integration for Hadoop allows IT organizations to gain visibility and manage dependencies between Hadoop and traditional jobs from one central location to meet SLAs.

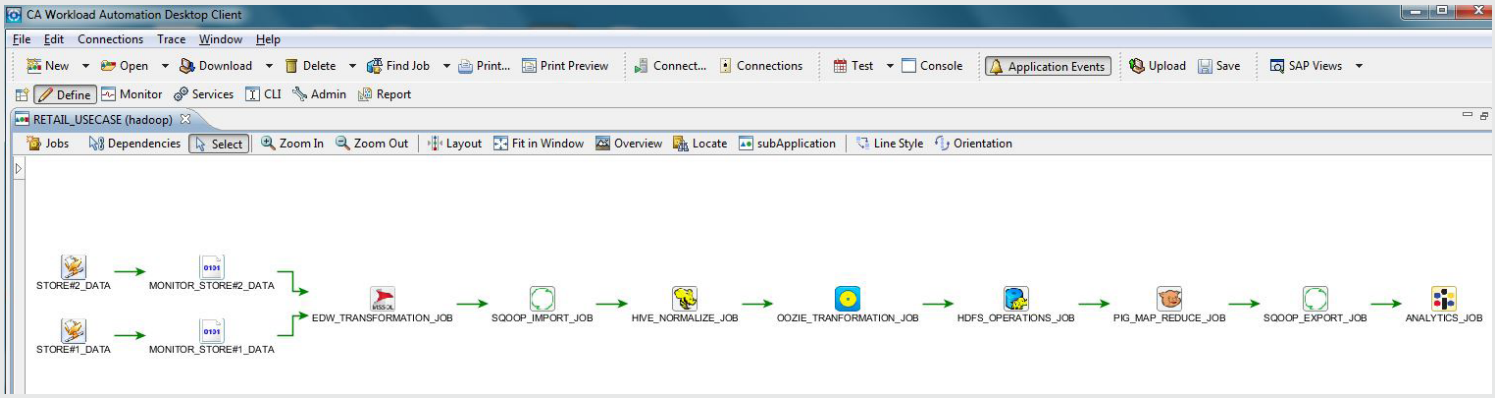
Key advantages of the solution include:

Seamless application integration: Extend scheduling capabilities and manage major business applications with reliability and flexibility and reduce Hadoop job complexity; execute in sync with workflows running in the rest of the enterprise.

Critical path analysis: Provides unique ability to identify and understand business impact of a Hadoop job within a job stream.

Multi-platform scheduling: Useres have the ability to manage and visualize a business process end-to-end across Hadoop and other platforms from a central point of control.

CA Workload Automation Advanced Integration for Hadoop lets you easily integrate, view and manage Hadoop jobs with from one central location.



Reduce cost and effort: Users have the ability to leverage CA Workload Automation jobs through familiar Hadoop interface, resulting in lower training costs. By removing the need to toggle between CA Workload Automation and Hadoop to manage dependencies, productivity is dramatically increased.

Improved ability to manage business processes: By integrating CA Workload Automation with Hadoop, business processes typically managed by separate groups can now be managed from within the same application, increasing visibility across the organization and enabling better IT and business practices alignment.

Critical Differentiators

- **Proven scalability and performance:** Scalability is obtained without running multiple instances, which is atypical compared to other scheduling engines.
- **Lowers cost:** This solution includes a enterprise-wide, single pane of glass that helps eliminate errors and reduces staff training.
- **Extreme flexibility:** Manage all workloads (distributed, mainframe and cloud) across all platforms in the environment, including Hadoop, Windows®, UNIX®, Linux®, HP NonStop, i5/OS™ and z/OS®.
- **No coding:** Intuitive GUI enables easy integration with Hadoop and other applications.
- **Rapid time-to-value:** Extend the use of stable, robust workload automation solution with CA Workload Automation Advanced Integration for Hadoop and enable quicker application deployment.
- **Built in fault tolerance and recovery:** CA Workload Automation provides reliable, lights-out operations with intelligent, automated recovery. This facilitates availability of business systems by enabling jobs to be completed accurately and on time.
- **Self-service:** This gives end-users the ability to request and execute workload processing controlled by workload policy and governance processes. This helps provide higher levels of efficiency and control while giving the business more control over their own services.

Related Products/Solutions

CA Workload Automation Agents enable organizations to automate workloads across multiple services and business applications, including SAP, PeopleSoft, Oracle E-Business, Informatica, Databases, MSSQL, Application Services, WebServices and Micro Focus.

CA Workload Automation iDash is designed to monitor the entire CA Workload Automation AE and CA Workload Automation CA 7® Edition environments, perform real-time forecasts based on the current status of the system, generate alerts when the thresholds are at risk of being missed and even execute automated recovery actions.

For more information, please visit **ca.com/wla**

CA Technologies (NASDAQ: CA) creates software that fuels transformation for companies and enables them to seize the opportunities of the application economy. Software is at the heart of every business, in every industry. From planning to development to management and security, CA is working with companies worldwide to change the way we live, transact and communicate – across mobile, private and public cloud, distributed and mainframe environments. Learn more at **ca.com**.

Copyright © 2015 CA. All rights reserved. Windows® is a registered trademark of Microsoft Corporation in the United States and/or other countries. UNIX is a registered trademark of The Open Group. Linux® is a registered trademark of the Linus Torvalds in the U. S. and other countries. i5/OS™ and z/OS® are trademarks of the International Business Machines Corporation in the United States, other countries, or both. All trademarks, trade names, service marks and logos referenced herein belong to their respective companies.

CS200-154207_0915