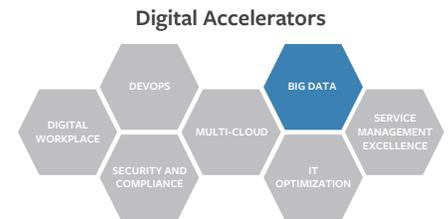


# Control-M for Hadoop

Simplify and automate Apache® Hadoop batch processing and connected enterprise workflows



## PRODUCT DESCRIPTION

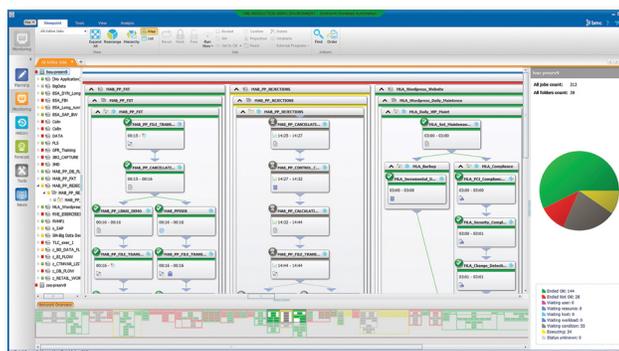
Control-M for Hadoop automates Hadoop batch processing and enables Hadoop workflows to be developed, scheduled, managed, and monitored with all other enterprise workloads in a single solution. It takes the complexity out of Hadoop management, accelerating implementation and delivering more accurate results.

## BUSINESS CHALLENGE

The data that Hadoop manages is often collected from different sources, such as enterprise data warehouses, relational databases, and traditional applications. These sources must be combined and synchronized to give Hadoop the timely input it needs. This integration is difficult to build, and development is typically done with scripts, which are time consuming to develop and maintain. **This complexity limits the speed at which Hadoop applications can be delivered, making it difficult to adapt to the challenges posed by increasingly diverse infrastructure, disparate data, and accelerated applications.**

## BMC SOLUTION

Control-M delivers the next wave of IT automation—**Digital Business Automation**—moving beyond traditional workload automation and adapting to modern IT technologies and processes. Control-M for Hadoop streamlines and simplifies Hadoop batch processing. It shortens development time by providing an intuitive interface for building workflows. It also includes native support for Hadoop, traditional platforms and applications, and file transfers, making it easier to create batch flows. **By simplifying workflow creation and management, Control-M for Hadoop helps enterprises get more out of their big data initiatives.**



## KEY FEATURES

Control-M for Hadoop simplifies workflow creation and management.

- **Familiar interface** – Manage Hadoop workloads through a simple drag-and-drop interface that provides universal visibility and control for all enterprise workloads
- **Flexible ecosystem support** – Apache® Spark®, Apache® Oozie®, Apache® DistCP®, Apache® Impala, and other jobs can all run through Control-M for Hadoop
- **Increased availability** – The co-location requirement for IBM® CICS®/IBM® DB2®, CICS/IBM® IMS® DB, IMS TM/DB2 is no longer an issue, allowing data access calls to be redirected upon a DB2 or IMS system failure and increasing availability
- **Improved reliability** – Detects and prevents potential problems with forecasting features and predictive analytics; restarts can be attempted automatically without requiring human intervention

## KEY BENEFITS

- **Makes implementation fast and accurate** by replacing manual scripting with automated workflow management and data integration, dramatically shortening development time, preventing coding errors, and reducing time-to-value
- **Improves business focus** instead of spending time building connectors and coordinating multiple tools

- The intuitive Control-M interface makes it easy to view, manage, and understand enterprise workloads.

## PRODUCT DETAILS

**Powerful, simple interface:** Users create and manage workflows via the interfaces they prefer. Graphical interfaces minimize the knowledge required of scheduling or Hadoop. Jobs are built and scheduled through wizard-like forms that include dropdown menus to simplify selection of the correct parameters and prevent input errors. Command line and programming interfaces enable developers to minimize scripting. Interfacing with enterprise applications, scheduling jobs, and monitoring their status are just as easy.

**Comprehensive support:** Control-M for Hadoop can manage workflows and integrate data from multiple sources. Supported solutions and formats include:

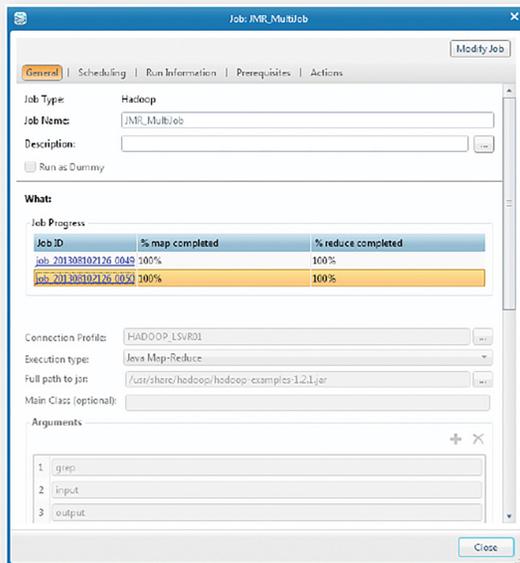
- **Hadoop** – Spark, Oozie, Impala, DistCp, Pig™, Hive™, Sqoop™, Tajo™, Streaming, Hive Server 2, and Hadoop Distributed File System (HDFS) operations
- **Enterprise Data Warehouses (EDW)** – Informatica®, IBM® InfoSphere® DataStage®, IBM DB2, Microsoft® SQL Server™, Oracle® databases, and others
- **Technologies** – FTP, sFTP, Web Services, Enterprise Java Beans (EJB), and Java® applications

- **ETL functions and business intelligence systems** – Informatica, IBM® Cognos®, SAP® BusinessObjects™, Oracle® Business Intelligence, and SQL Server SSIS
- **RDBMS** – Oracle, Sybase®, DB2/UDB, SQL Server, and PostgreSQL®
- **Virtualization and cloud** – Amazon® EC2™, BladeLogic, and VMware®
- **Backup systems** – Symantec® NetBackup™ and IBM® Tivoli® Storage Manager
- **Operating systems** – Microsoft® Windows®, UNIX®, Linux®, IBM® z/OS®, IBM® iSeries®, Tandem®, and OpenVMS

**Self-service:** Users can monitor and manage their own workloads from any browser-based device, including mobile iOS® and Android™ devices.

## FOR MORE INFORMATION

To learn more about Control-M for Hadoop, please visit [bmc.com/hadoop](http://bmc.com/hadoop)



- Forms and wizards take the place of scripting to develop workflows, even those that require complex coordination among multiple data sources and enterprise systems.

Certified Hadoop partner with IBM, Cloudera, Hortonworks, and MapR



BMC is a global leader in innovative software solutions that enable businesses to transform into digital enterprises for the ultimate competitive advantage. Our Digital Enterprise Management solutions are designed to fast track digital business from mainframe to mobile to cloud and beyond.

**BMC – Bring IT to Life**

**BMC digital IT transforms 82 percent of the Fortune 500.**



BMC, BMC Software, the BMC logo, and the BMC Software logo, and all other BMC Software product and service names are owned by BMC Software, Inc. and are registered or pending registration in the US Patent and Trademark Office or in the trademark offices of other countries. All other trademarks belong to their respective companies. © Copyright 2017 BMC Software, Inc.

