

CA Unified Infrastructure Management for Hadoop



At a Glance

When leveraging Hadoop®, organizations can realize a wealth of insights, but they can also encounter a lot of administrative complexity. With CA Unified Infrastructure Management (CA UIM, formerly CA Nimsoft Monitor), enterprises and service providers gain the comprehensive monitoring coverage they need to efficiently track the entire Hadoop ecosystem and proactively optimize its performance. CA UIM provides availability and performance metrics on all the key elements of the Hadoop framework, including clusters, master daemons and Hadoop Distributed File System™ (HDFS). Plus, it provides a centralized interface for monitoring the infrastructure that supports Hadoop, such as network devices, operating systems and more.

Key Benefits/Results

- **Unified visibility.** Incorporate Hadoop monitoring into a central platform that can track all the environments that matter to your business.
- **Improve service levels.** Leverage insights and alerts that help your team detect and resolve issues faster, reduce outage duration and address issues before service levels take a hit.
- **Reduce cost and complexity.** Harness the comprehensive coverage and capabilities that can reduce the costs associated with buying and supporting multiple, disconnected and complex tools.

Key Features

- **Comprehensive coverage.** Get performance metrics on all the key elements of the Hadoop framework and any supporting infrastructure.
- **Powerful, timely insights.** Get automated alerts, trending reports and real-time intelligence that offer the insights you need, when you need them.
- **Customizable reports and real-time dashboards.** Adapt information views to specific roles and objectives, and leverage historical performance reports and SLA reports.

Business Challenges

Perhaps more than any single technology, Hadoop has helped bring big data capabilities to the masses, offering organizations an open-source software framework for distributing storage and processing across clusters of commodity hardware. Through these capabilities, organizations can economically employ big data techniques to mine more value out of their information repositories.

However, for those tasked with managing Hadoop environments and service levels, the very advantages these technologies provide, namely highly distributed, processing-intensive workloads, can pose significant challenges. How do you track performance across all the nodes in operation, as well as the complex ecosystem of hardware and networking equipment these systems rely on? With so many inter-reliant components, how do you quickly determine whether service levels are on target, and, if not, where the issue is?

Solution Overview

CA UIM represents a comprehensive IT management platform that provides the capabilities administrators need to track and optimize the performance of their Hadoop environments. Further, the product offers the combination of centralized visibility, ease of deployment and operation, and architectural efficiency that can significantly reduce the cost and effort associated with Hadoop administration.

Without CA UIM, monitoring Hadoop can be a time consuming and complex activity. Depending on the scope of the deployment, dozens or hundreds of clusters can be operating simultaneously. Further, even if a point solution has been employed to track all these elements, still other tools would be needed to track the associated infrastructure elements. Beyond the up-front cost and effort of implementation, these tools require substantial ongoing effort associated with manual data aggregation, analysis and reporting. With CA UIM, you can eliminate or substantially reduce all this cost and effort. The platform itself is easy to deploy and manage, and it enables you to automatically generate the views and insights you need to manage the environment.

CA UIM equips you with a single, unified platform that enables you to monitor and control your entire IT environment. With the platform you can gain timely alerts and insights that help speed issue detection and resolution, and help spot ways to pre-empt issues and institute tactics for ongoing optimization.

Critical Differentiators

Comprehensive coverage of Hadoop and beyond

To effectively track Hadoop performance, your administrators need to be able gain visibility into all the elements that make up the environment, including nodes, clusters, master daemons and HDFS. With CA UIM, you can uniformly track these complex and dynamic environments. Further, CA UIM provides out-of-the-box-support for more than 140 device types—including servers, databases, applications, network devices and cloud services. As a result, you can leverage your monitoring investment to centrally track all the elements that support the Hadoop ecosystem—and the rest of your business.

Efficient, resilient architecture

The CA UIM architecture is based on a unique message bus that uses a publish/subscribe framework. This architecture inherently enables scalability, high availability, extensibility and rapid deployment. By standardizing the communication between data collection, data aggregation and data analysis, the architecture provides a single and extensible implementation across all computing components and infrastructures. With CA UIM, there is no stitching together of multiple products and solutions as with other alternatives.

Optimal ease of use and efficiency

CA UIM offers the speed and ease of use associated with point solutions, featuring a unique architecture with a lightweight

footprint. At the same time, the solution delivers enterprise scalability and multi-tenancy. Further, through its extensible monitoring coverage, CA UIM enables customers to add monitoring of new technologies and environments, without having to make new product, staffing or training investments. With CA UIM, IT operations can reduce the time and cost associated with monitoring, while business and IT management can gain the insights needed to optimize resources, planning and investments.

Timely, powerful insights

CA UIM offers all the capabilities you need to ensure that your team can gain maximum insights from the monitoring data captured:

- **Alarm management.** CA UIM helps IT organizations reduce the time it takes to troubleshoot problems by consolidating events from various devices, servers and applications and providing the ability to automatically assign alarms to the right IT staff.
- **Dashboards.** CA UIM dashboards provide real-time views of IT and business services. Dashboards can be tailored to specific users, groups and roles, enabling users to see practically any type of data in a single view—including help desk call statistics, application performance metrics, IT resource utilization and much more.
- **Reporting.** With CA UIM, you can easily develop reports based on any monitoring data captured by the solution. This includes not only Hadoop elements but

internal servers and networking equipment, virtualized infrastructures, SaaS offerings, cloud-based services and outsourced environments.

Granular, extensive coverage of Hadoop environments

CA UIM offers availability, health check, and basic up/down tracking of all master daemons in Hadoop environments, including NameNode, Resource Manager and HBaseMaster. The solution offers coverage of metrics in these areas:

- **Clusters:** Throughput, storage capacity utilization as a %, %DataNodes down, %DataNodes storage full.
- **Resource Manager workload characteristics:** Wait times, run times, maps/reducers per job, input/output/shuffle data sizes per job, etc.
- **NameNode:** Operation counts, heartbeats, JVM garbage collection (ms), heap memory used, uptime, RPC latency/average wait time, edit log directories, write errors.
- **SecondaryNameNode/DataNode:** Live, dead or decommissioned; capacity.
- **HDFS:** Storage capacity, block errors (corrupt, missing, etc.), thread state, thread contention stats, under replicated blocks, blocks pending replication.
- **HBase™:** Cluster requests/hr, average load, master heap information, garbage collection stats, % region servers down.
- **RegionServer:** Read requests, write requests, regions served, average read latency, average write latency, JVM garbage collection, average waiting time.

For more information, please visit ca.com/uim

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.