SOLUTION BRIEF

Securing Sensitive Data in Hadoop Clusters with Gemalto SafeNet ProtectFile

The Problem

From large enterprises to start-ups and small businesses, companies of every size are generating more data than ever before. In fact, IDC recently reported the global volume of data will grow by a factor of 300—from 130 to 40,000 exabytes—between the years 2005 to 2020[1]. To gain valuable insights and intelligence from big data, enterprises need a powerful framework that can process and store growing volumes of data from a variety of sources, and in a variety of formats. One emerging technology that addresses this need is Apache Hadoop, an open source framework that distributes the processing of large data sets across clusters on low-cost hardware.

Organizations deploy Hadoop because of the benefits it can provide, including scalable, costeffective storage and faster data processing. As a result of these efficiency gains, enterprises increase their reliance on Hadoop as their corporate data asset grows. Often, it's not until well after Hadoop is deployed that a company realizes a very important element that is missing from its big data implementation – the ability to secure all of the sensitive data now distributed across DataNodes in clusters.

Hadoop stores information in clusters that it distributes across hundreds, and sometimes thousands, of DataNodes. Yet, it lacks the ability to completely secure the data-at-rest residing in those clusters. Each of these nodes represents a potential entry point for a rogue insider or malicious threat. If an unauthorized user or service assumes direct access to a node, sensitive data is in clear view. This presents a tremendous, and potentially costly, risk for organizations. Indeed, the inability to lock down and protect data is one factor that will slow the continued adoption of Hadoop – particularly among enterprises that must meet strict compliance and regulatory mandates.

SafeNet ProtectFile for Apache Hadoop Highlights

Seamless encryption of big data implementations

> Provides transparent and automated encryption of sensitive data in clusters with minimal impact on Hadoop performance or end-user experience.

Rapid deployment and implementation

> Utilize automation tools for fast, easy roll-out and standard deployment to multiple DataNodes in a Hadoop cluster.

No rearchitecting required

> No changes to your existing big data implementation are necessary.

Centralized key management

> Manage risk and maintain control of encryption keys for robust security and scalability.

Granular access controls

> Define and enforce policies to guard against unauthorized and rogue access to, and possible exposure of, high value data.

Maintain compliance

> Support compliance mandates, such as HIPAA and PCI DSS, in your big data implementation.

The Solution

Today, insider threats are prevalent and security breaches are inevitable. Because encryption is not available in Hadoop, enterprises deploying the popular framework to store and process big data must take steps to secure high-value data now.

SafeNet ProtectFile provides transparent and seamless encryption of sensitive data stored in Hadoop clusters with minimal impact on performance or the end-user experience. As a complete, enterprise-ready encryption solution, SafeNet ProtectFile provides automation tools for fast and easy deployment across large numbers of DataNodes in a Hadoop cluster, as well as large-scale deployment of SafeNet

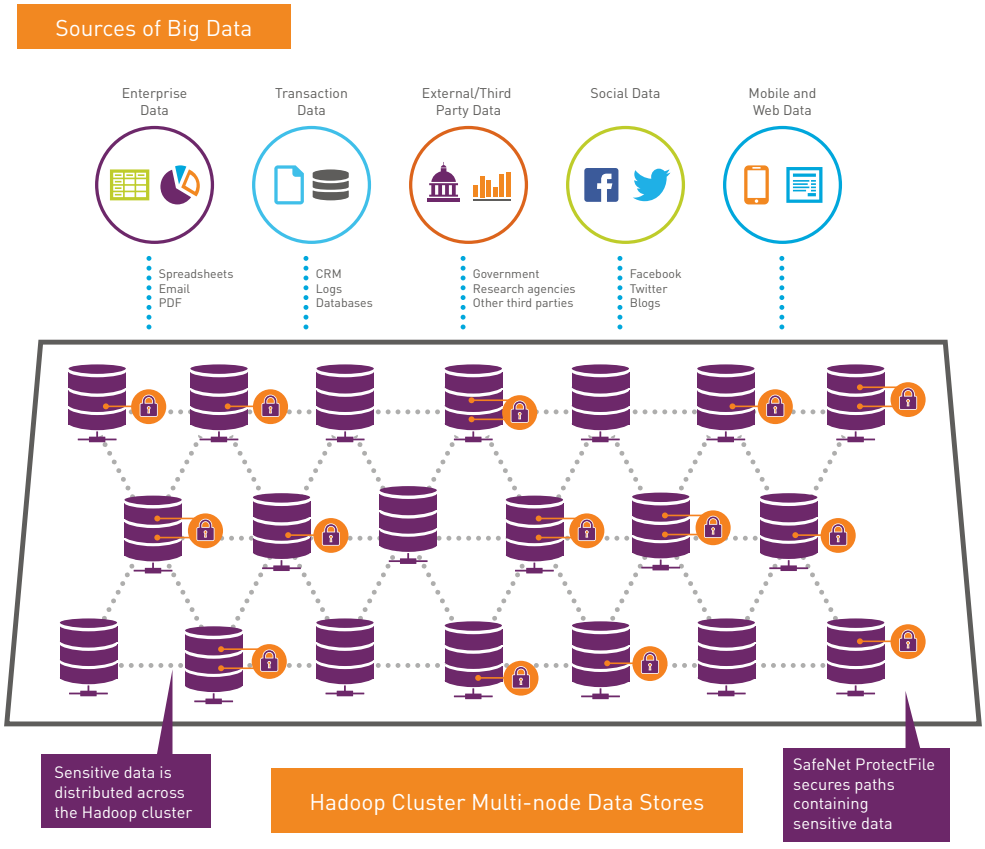
1. http://idc-cema.com/eng/events/54146-idc-big-data-and-business-analytics-forum-2014

ProtectFile Linux across multiple servers. In addition, no changes to an enterprise's existing big data implementation are necessary.

Once SafeNet ProtectFile is deployed, Hadoop services continue to run as expected, while encryption and decryption of data happens transparently to Hadoop services and end users. When encryption is applied to sensitive information on DataNodes, it will always remain secure while at rest on disk. Any attempt to access the encrypted data by an unauthorized service or process will be blocked.

To manage risk and maintain control, SafeNet ProtectFile works in tandem with Gemalto's FIPS 140-2 Level 3-certified SafeNet KeySecure hardware appliance for centralized key and policy management. SafeNet ProtectFile stores keys separately from protected data to guard against unauthorized access to high-value information stored in Hadoop clusters. Granular access controls enable the creation and enforcement of policies to define users and Hadoop services authorized to access the encrypted data.

For enterprises with strict organizational or regulatory mandates, such as HIPAA and PCI DSS, SafeNet ProtectFile makes big data implementations with Hadoop possible. Strong encryption and advanced key management, as well as comprehensive auditing and reporting capabilities, ensure organizations meet compliance requirements while protecting sensitive data in Hadoop clusters against malicious attacks.



Request More Information

If your enterprise has deployed or is planning to deploy Hadoop for big data management, you need to have a plan to protect your data. Contact us today to get started with SafeNet ProtectFile for transparent and seamless encryption of your sensitive data stored in Hadoop clusters – all without disrupting performance or end-user experience.

Contact Us: For all office locations and contact information, please visit www.safenet-inc.com
Follow Us: data-protection.safenet-inc.com

GEMALTO.COM

gemalto
security to be free

©2015 SafeNet, Inc. All rights reserved. SafeNet and SafeNet logo are registered trademarks of SafeNet. All other product names are trademarks of their respective owners. SB (EN)-Apr.02.2015 - Design: Jubemo