A Dataguise Case Study

SAMSUNG

The Challenge

Performing product analytics on millions of Samsung Galaxy Smartphone devices worldwide while ensuring personal private information stays protected.

Dataguise Solution

DgSecure for Hadoop automatically discovers consumer privacy data, device and encrypts it before hitting the cloud in 7 Amazon AWS clusters globally.

Case Study Highlights

Global Data Protection For PII Data

- On-the-fly Flume protection
- Locking only names, Device Ids
- Non-blocking to analytics deployments

100% Flexible to Samsung's Requirements

- Drop-in solution (no coding)
- Functions across AWS EMR, S3, Hortonworks, Pivotal HD, Files
- High availability (<1min recovery)

Global Leader in Product Analytics

Samsung has been analyzing and improving mobile and smart tv products through product analytics for decades. Through that period, a number of different tools, approaches, data repositories, data capture, and data storage locations have been employed. To improve product performance, reliability, feature adoption, ease of use, Samsung has captured realms of device-specific data, such as the location, hardware specifications, utilization rates, capacity and battery life, etc. Hadoop makes the processing, collection, analytics of this data faster and move cost effective for Samsung.

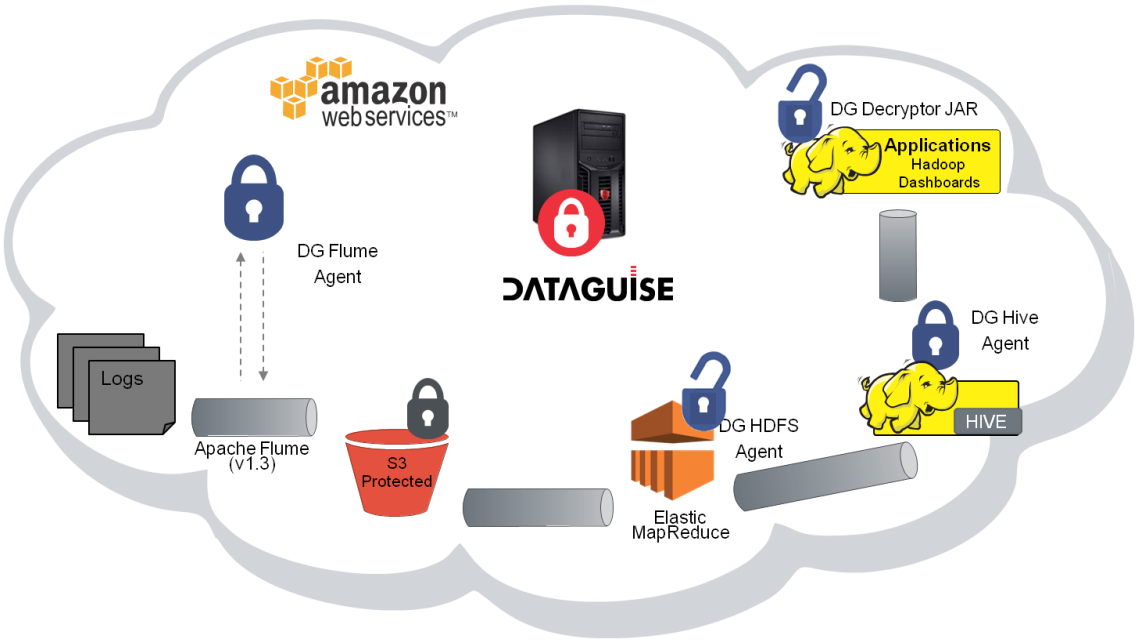
The Landscape Has Changed

As a global manufacturer with products in all markets and territories, Samsung also must adequately protect any sensitive data from device logs and data capture. Specifically, in Europe, new privacy policies defined in the European Union Privacy Directive require Samsung to protect any personal identifiable information specific to European citizens. Samsung still needed to collect device data for analytics, but was mindful of privacy laws, and privacy fines levied on competitors that did not fully comply with privacy mandates.

Big Data Protection Goals

- Aggregate logging data (product, usage, user configuration) for all smartphones worldwide
- De-identify personal user info to ensure privacy and compliance with European/US Privacy
- Keep all sensitive data encrypted at-rest, and provide authorized access (decryption) of sensitive data on a case-by-case basis for analytics applications that require access to full, complete, plaintext data.

The Solution Enviroment



The Solution

- Samsung utilizes Dataguise Flume agent to protect all sensitive data being written to Amazon S3
- Runs Dataguise in AWS, also utilizes Dataguise EMR security agents to selectively decrypt for
- authorized analytics in AWS
- Achieves On-demand Hadoop for product analytics, user behavior, supply chain optimization in high scale-out, high performance and high availability system
- 100% cloud based

About Dataguise

Dataguise is the leading provider of data-centric security and data governance solutions for Big Data. Organizations in financial services, healthcare, retail, government and other industries that value data privacy and are subject to regulatory compliance rely on Dataguise for discovery, protection (encryption, masking, redaction), audit and intelligence for sensitive data in Hadoop, RDBMSs and other Big Data environments. Dataguise enables businesses to balance the use of Big Data to drive business while protecting privacy and personal identifiable information (PII), maintaining compliance and minimizing the risks of exposure of legally protected or regulated data covered by PCI-DSS, HIPAA, HITECH and the like.

DATAGUISE

2201 Walnut Ave, Suite 260
Fremont, CA 94538
P: +1 510.824.1036
F: +1 866.384.8082

www.dataguise.com