Securing data lakes and analytics on AWS

Nikki Rouda, Principal PMM
March 28th, 2019
Bringing it all together

Data Subjects

Customers are Controllers

AWS as Processor

Controllers and Processors have obligations under GDPR
Bringing it all together

Controllers and Processors have obligations under GDPR

Data Subjects
Customer's customer as Controller
Customer as Processor
AWS as Processor
**GDPR in practice: implementing TOMs**

Under GDPR, controllers and processors are required to implement appropriate technical and organizational measures (TOMs) ...

<table>
<thead>
<tr>
<th>(1) Pseudonymization and encryption of personal data</th>
<th>(2) Ensure ongoing confidentiality, integrity, availability, and resilience of processing systems and services</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3) Ability to restore availability and access to personal data in a timely manner in the event of a physical or technical incident</td>
<td>(4) Process for regularly testing, assessing, and evaluating the effectiveness of TOMs</td>
</tr>
</tbody>
</table>
Today, I'm very pleased to announce that AWS services comply with the General Data Protection Regulation (GDPR). This means that, in addition to benefiting from all of the measures that AWS already takes to maintain services security, customers can deploy AWS services as a key part of their GDPR compliance plans.
What AWS provides

Tools and services

Compliance framework

AWS Marketplace

Partner network

Data protection terms
AWS shared responsibility model

Customers

- Customer content
- Platform, Applications, Identity & Access Management
- Operating System, Network & Firewall Configuration
- Client-side Data Encryption
- Server-side Data Encryption
- Network Traffic Protection

AWS Foundation Services

- Compute
- Storage
- Database
- Networking

AWS Global Infrastructure

- Availability Zones
- Regions
- Edge Locations

Customers are responsible for their security and compliance IN the cloud.

AWS is responsible for the security OF the cloud.
Compliance is also a “shared responsibility”

<table>
<thead>
<tr>
<th>Feature</th>
<th>Controllers responsibility</th>
<th>Processors responsibility</th>
<th>AWS has tools to help</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal Compliance</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>System Security and Data Protection by Design</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Records of Processing Activities</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Encryption</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Managing Data Subject Consent</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Managing Personal Data Deletion</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Managing Personal Data Portability</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Security of Personal Data</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The controller “shall implement appropriate technical and organizational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed.”

Multi-factor authentication
API-Request Authentication
Temporary Access Tokens
“Each controller and, where applicable, the controller’s representative, shall maintain a record of processing activities under its responsibility.”

CloudTrail
Amazon Inspector
Macie
AWS Config
Compliance tools

Organizations must “implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk, including the pseudonymization and encryption of personal data.”

Encryption of your data at rest with AES256 (EBS/S3/Amazon Glacier/RDS)
Centralized (by Region) with Key Management (AWS KMS)
IPsec tunnels into AWS with the VPN-Gateways
Dedicated HSM modules in the cloud with CloudHSM
Appropriate technical and organizational measures may need to include “the ability to ensure the ongoing confidentiality, integrity, availability, and resilience of the processing systems and services.”

SOC 1 / SSAE 16 / ISAE 3402 (formerly SAS 70) / SOC 2 / SOC 3
PCI DSS Level 1
ISO 9001 / ISO 27001 / ISO 27017 / ISO 27018
FIPS 140-2
C5
Meet your own security objectives

Customers
- Your own accreditation
  - GDPR Code of Conduct
- Your own certifications
- Your own external audits

Customer scope and effort is reduced
Better results through focused efforts

AWS Foundation Services
Built on AWS consistent baseline controls

AWS Global Infrastructure
GDPR center

https://aws.amazon.com/compliance/gdpr-center/
Most services for security and governance

Customers need to have multiple levels of security, identity and access management, encryption, and compliance to secure their data lake

<table>
<thead>
<tr>
<th>Security</th>
<th>Identity</th>
<th>Encryption</th>
<th>Compliance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon GuardDuty</td>
<td>AWS IAM</td>
<td>AWS Certification Manager</td>
<td>AWS Artifact</td>
</tr>
<tr>
<td>AWS Shield</td>
<td>AWS SSO</td>
<td>AWS Key Management Service</td>
<td>Amazon Inspector</td>
</tr>
<tr>
<td>AWS WAF</td>
<td>Amazon Cloud Directory</td>
<td>Encryption at rest</td>
<td>Amazon Cloud HSM</td>
</tr>
<tr>
<td>Amazon Macie</td>
<td>AWS Directory Service</td>
<td>Encryption in transit</td>
<td>Amazon Cognito</td>
</tr>
<tr>
<td>VPC</td>
<td>AWS Organizations</td>
<td>Bring your own keys, HSM support</td>
<td>AWS CloudTrail</td>
</tr>
</tbody>
</table>
### Most secure — certifications

<table>
<thead>
<tr>
<th>Global</th>
<th>United States</th>
<th>European</th>
<th>Asian Pacific</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSA Cloud Security Alliance Controls</td>
<td>CJS Criminal Justice Information Services</td>
<td>ITAR International Arms Regulations</td>
<td></td>
</tr>
<tr>
<td>ISO 9001 Global Quality Standard</td>
<td>DoD SRG DoD Data Processing</td>
<td>MPAA Protected Media Content</td>
<td></td>
</tr>
<tr>
<td>ISO 27001 Security Management Controls</td>
<td>FedRAMP Government Data Standards</td>
<td>NIST National Institute of Standards and Technology</td>
<td></td>
</tr>
<tr>
<td>ISO 27017 Cloud Specific Controls</td>
<td>FERPA Educational Privacy Act</td>
<td>SEC Rule 17a-4(f) Financial Data Standards</td>
<td></td>
</tr>
<tr>
<td>ISO 27018 Personal Data Protection</td>
<td>ISO FFIEC Financial Institutions Regulation</td>
<td>VPAT/Section 508 Accountability Standards</td>
<td></td>
</tr>
<tr>
<td>SOC 3 General Controls Report</td>
<td>HIPAA Protected Health Information</td>
<td>K-ISMS [Korea] Korean Information Security</td>
<td></td>
</tr>
</tbody>
</table>

© 2019, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential
Challenges to making a secure data lake

Typical steps of building a data lake

1. Setup storage
2. Move data
3. Cleanse, prep, and catalog data
4. Configure and enforce security and compliance policies
5. Make data available for analytics
Build a secure data lake in days with AWS Lake Formation

- Move, store, catalog, and clean your data faster
- Enforce security policies across multiple services
- Gain and manage new insights

Move, store, catalog, and clean your data faster with Machine Learning

Enforce security policies across multiple services

Empower analyst and data scientist to gain and manage new insights
The Forrester Wave

Cloud Hadoop/Spark Platforms
Q1 2019

The 11 Providers That Matter Most and How They Stack Up

by Noel Yuhanna and Mike Gualtieri
February 13, 2019

The Forrester Wave™ is copyrighted by Forrester Research, Inc. Forrester and Forrester Wave™ are trademarks of Forrester Research, Inc. The Forrester Wave™ is a graphical representation of Forrester's call on a market and is plotted using a detailed spreadsheet with exposed scores, weightings, and comments. Forrester does not endorse any vendor, product, or service depicted in the Forrester Wave™. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change.
Thank you!
nrouda@