Big Data Innovation Conference BDIC2016 Workshop

Evaluating Security and Compliance of the Outsourced Cloud based Big Data Infrastructures

Yuri Demchenko
System and Network Engineering, University of Amsterdam

BDIC2016 Conference
16 November 2016, Frankfurt
Outline

Part 1. Big Data and Cloud Security
• Big Data Architecture Framework (BDAF)
• Big Data Infrastructure and Applications Stack
• Cloud security models, services and mechanisms
• Cloud Security best practices: AWS and Microsoft Azure

Part 2. Cloud Compliance and (Self-) Assessment
• Compliance standards, Security Controls
• CSA GRC Stack: Governance, Risk Management and Compliance
• PCI DSS Cloud Computing Guidelines

Part 3. Privacy protection and regulations
• Privacy related standards
• EU General Data Protection Regulation (GDPR)

Part 4. Exercise/Hands-on
• CSA Consensus Assessment Initiative Questionnaire
• Exercise/Hands-on: PCI DSS Self-Assessment Questionnaire
Cloud Computing is the New Pervasive Ubiquitous Computing, Communications and Data Storage Platform

Day to Day Life

Education

Relationships

Entertainment

Communications

Infrastructure

Transportation

Commerce

Slide is credit to David Bernstein
Public Clouds are All Over the Places

Cloud Centers are All Over the Place, from datacentermap.com November 2016
http://www.datacentermap.com/cloud.html
Big Data is Driving Cloud Usage

There were 5 exabytes of information created between the dawn of civilization through 2003, but that much information is now created every 2 days, and the pace is increasing.
Eric Schmidt, Google CEO, Techonomy Conference, August 4, 2010

Data is becoming the new raw material of business: an economic input almost on a par with capital and labour. “Every day I wake up and ask, ‘how can I flow data better, manage data better, analyse data better?’” says Rollin Ford, the CIO of Wal-Mart.
Source: Data, Data Everywhere, The Economist, February 25, 2010

- Built on experience of developing the Cloud Computing standards fully accepted by industry


- Volume 1: NIST Big Data Definitions
- Volume 2: NIST Big Data Taxonomies
- Volume 3: NIST Big Data Use Case & Requirements
- Volume 4: NIST Big Data Security and Privacy Requirements
- Volume 5: NIST Big Data Architectures White Paper Survey
- Volume 6: NIST Big Data Reference Architecture
- Volume 7: NIST Big Data Technology Roadmap

NBD-WG defined 3 main components of the new technology:

- Big Data Paradigm
- Big Data Science and Data Scientist as a new profession
- Big Data Architecture

The **Big Data Paradigm** consists of the distribution of data systems across horizontally-coupled independent resources to achieve the scalability needed for the efficient processing of extensive datasets.
Main components of the Big Data ecosystem
- Data Provider
- Big Data Applications Provider
- Big Data Framework Provider
- Data Consumer
- Service Orchestrator

Big Data Lifecycle and Applications Provider activities
- Collection
- Preparation
- Analysis and Analytics
- Visualization
- Access

Big Data Ecosystem includes all components that are involved into Big Data production, processing, delivery, and consuming

Big Data Architecture Framework (BDAF)

(1) Data Models, Structures, Types
   – Data formats, non/relational, file systems, etc.

(2) Big Data Management
   – Big Data Lifecycle (Management) Model
     • Big Data transformation/staging
   – Provenance, Curation, Archiving

(3) Big Data Analytics and Tools
   – Big Data Applications
     • Target use, presentation, visualisation

(4) Big Data Infrastructure (BDI)
   – Storage, Compute, (High Performance Computing,) Network
   – Sensor network, target/actionable devices
   – Big Data Operational support

(5) Big Data Security
   – Data security in-rest, in-move, trusted processing environments
Big Data Infrastructure and Analytics Tools

Big Data Infrastructure
- Heterogeneous multi-provider inter-cloud infrastructure
- Data management infrastructure
- Collaborative Environment
- Advanced high performance (programmable) network
- Security infrastructure
- Federated Access and Delivery Infrastructure (FADI)

Big Data Analytics Infrastructure/Tools
- High Performance Computer Clusters (HPCC)
- Big Data storage and databases SQL and NoSQL
- Analytics/processing: Real-time, Interactive, Batch, Streaming
- Big Data Analytics tools and applications
Data Lifecycle/Transformation Model

- Data Model changes along data lifecycle or evolution
- Data provenance is a discipline to track all data transformations along lifecycle

- Identifying and linking data
  - Persistent data/object identifiers (PID/OID)
  - Traceability vs Opacity
  - Referral integrity

Multiple Data Models and structures
- Data Variety and Variability
- Semantic Interoperability

Data (inter)linking
- PID/OID
- Identification
- Privacy, Opacity
- Traceability vs Opacity

Data Storage (Big Data capable)
The major structural components of the Big Data stack are grouped around the main stages of data transformation

- **Data ingest**: Ingestion will transform, normalize, distribute and integrate to one or more of the Analytic or Decision Support engines; ingest can be done via ingest API or connecting existing queues that can be effectively used for handles partitioning, replication, prioritisation and ordering of data

- **Data processing**: Use one or more analytics or decision support engines to accomplish specific task related to data processing workflow; using batch data processing, streaming analytics, or real-time decision support

- **Data Export**: Export will transform, normalize, distribute and integrate output data to one or more Data Warehouse or Storage platforms;

- **Back-end data management, reporting, visualization**: will support data storage and historical analysis; OLAP platforms/engines will support data acquisition and further use for Business Intelligence and historical analysis.
Big Data Stack

Hook into an existing **queue** to get copy / subset of data. Queue handles partitioning, replication, and ordering of data, can manage backpressure from slower downstream components.

**Ingestion** will transform, normalize, distribute and integrate to one or more of the Analytic / Decision Engines of choice.

Use one or more Analytic / Decision engines to accomplish specific task on the Fast Data window.

**OLAP Engines to support data acquisition for later historical data analysis and Business Intelligence on entire Big Data set**.

**Real-Time Decisions** and/or Results can be fed back “up-stream” to influence the “next step”.

**Export** will transform, normalize, distribute and integrate to one or more Data Warehouse or Storage Platforms.

**Use direct Ingestion API to capture entire stream at wire speed**.
Big Data Stack

Hook into an existing **queue** to get copy / subset of data. Queue handles partitioning, replication, and ordering of data from slower downstream components.

**Ingestion** will transform, normalize, distribute and integrate to one or more of the Analytic / Decision Engines of choice.

Use one or more Analytic / Decision engines to accomplish specific task on the Fast Data window.

OLAP Engines to support data acquisition for later historical data analysis and Business Intelligence on entire Big Data set.

**Real-Time Decisions** and/or Results can be fed back “up-stream” to influence the “next step”.

**Ingestion API** to capture entire stream at wire speed.

**Ingestion** will transform, normalize, distribute and integrate to one or more of the Analytic / Decision Engines of choice.

**Business Process Management** (for biz competences)

**Export** will transform, normalize, distribute and integrate to one or more Data Warehouse or Storage Platforms.

**Data Ingestion**
- Message/Data Queue
- Use one or more Analytic / Decision engines to accomplish specific task on the Fast Data window
- Ingestion will transform, normalize, distribute and integrate to one or more of the Analytic / Decision Engines of choice
- OLAP Engines to support data acquisition for later historical data analysis and Business Intelligence on entire Big Data set

**Data Export**
- Use direct Ingestion API to transform, normalize, distribute and integrate to one or more Big Data Engines to support data acquisition for later historical data analysis and Business Intelligence on entire Big Data set

**Data Management**
- Design Experiment
- Test Hypothesis
- Analyse Data
- Identify Patterns
- Scientific Methods

**Real-Time Decisions**
- Continuously query
- OLTP Model
- Prog Request - Response
- Stored Procedures
- Export for Pipelining
- Business Intelligence, Reports
- User Interactive

**Batch Processing**
- Search Query
- Decision support
- Statistics
Cloud Platform Benefits for Big Data

- **Cloud deployment on virtual machines or containers**
  - Applications portability and platform independence, on-demand provisioning
  - Dynamic resource allocation, load balancing and elasticity for tasks and processes with variable load

- **Availability of rich cloud based monitoring tools for collecting performance information and applications optimisation**

- **Network traffic segregated and isolation**
  - Big Data applications benefit from cloud based clusters: dynamic cluster resizing, load balancing, and other scale-out operations
  - Internal clouds network separates networks traffic for data and for management
  - Layer 2 and Layer 3 virtual networks inside user/application VPC

- **Cloud tools for large scale applications deployment and automation**
  - Provide basis for agile services development and Zero-touch services provisioning
  - Applications deployment in cloud is supported by major Integrated Development Environment (IDE)
  - Built-in interfaces for content distribution and mobile access
Cloud Based Big Data Services

Characteristics:
Massive data and computation on cloud, small queries and results

Examples:
Search, scene completion service, log processing
Cloud HPC and Big Data Platforms

• HPC on cloud platform
  – Special HPC and GPU VM instances as well as Hadoop/HPC clusters offered by all CSPs

• Amazon Big Data services
  – Amazon Elastic MapReduce, Kinesis, DynamoDB, Regshift, etc

• Microsoft Analytics Platform System (APS)
  – Microsoft HD Insight/Hadoop ecosystems

• IBM BlueMix applications development platform and Watson cognitive platform
  – Includes full cloud services and advanced data analytics services

• LexisNexis HPC Cluster System
  – Cyber Analytics platform and Government Data Analysis

• Variety of Open Source tools
  – Streaming analytics/processing tools: Apache Kafka, Apache Storm, Apache Spark
Responsibilities Split in IaaS, PaaS, SaaS

Security management responsibilities split between Customer and Provider for IaaS, PaaS, SaaS service models:

- Updating firmware and software for platform and for customer managed components
- Firewall is intrusion prevention is a responsibility of the cloud provider
- Certification and compliance of the cloud platform doesn't imply security and compliance of the customer controlled components
### Data Protection Obligations by CSP and Customers

- Although customers are responsible for classifying their data, cloud providers should make written commitments to customers about the privacy of the customer data stored within their cloud.
- These commitments should include information about privacy and security practices, data use limitations, and regulatory compliance.
- Cloud providers should make certifications and audit reports that demonstrate compliance with key standards and regulations.
- Customers should not migrate data to a cloud provider that cannot address their data protection needs.

<table>
<thead>
<tr>
<th>Responsibility</th>
<th>IaaS</th>
<th>PaaS</th>
<th>SaaS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data classification and accountability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Client and end point protection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identity and access management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application level controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host security</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical security</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Blue = Cloud customer
- Black = Cloud provider
Security Basics: What should you know about Security?

• **Password is a basis for secure access** but it is not enough to secure your applications and services.
  – There is whole stack of network and infrastructure or platform security services and mechanisms which need to be applied in a consistent way to ensure high system **dependability** and **availability**

• **Basis for secure communication and data transfer** are the security protocols and security mechanisms
  – Security services are defined for communicating entities and can work at different layers
  – Security mechanisms can be used by services and functional components to achieve one or another aspect of security

• **Security is an overloaded term and may mean different aspects**
  - **Network/communication Security** - **Data Security** – **Application Security** - **Operation Security** – **System Security**

• **What kind of data to protect**
  – **Application Data** – **Personal Data (User ID, personal information)** – **Infrastructure management data**

• **Data security must be considered for at least 3 aspects**
  - **Data in transfer (Communication)** – **Data in-rest (Stored)** – **Data at run-time (Processed)**

• **Relations between Security and Trust**
  – Trust or trust relations is a foundation of the security protocols and services
  – Technical trust is typically based on possession of shared secret or private key in Public Key Cryptography
  – Social trust may include aspects such as reputation and may have different measures for level of trust
Different sides of Security and Trust

- Modern paradigm of remote distributed services and online/downloadable digital content provisioning makes security and trust relations between User and Provider more complex.
- User and Service Provider – the two actors concerned with own Data/Content security and each other System/Platform trustworthiness.
- Two other aspects of security/trust:
  - Data stored vs Data processed
  - System Idle vs Active (running User session)
Trust Relations Between Provider and Customer

Components of the trust relations when processing data on the provider platform

- Data stored on user facilities
- Data processed and stored on provider facilities
- User client must trust provider’s system/platform
- Provider knows/trusts user by their ID credentials (e.g. username/password)
Cloud, OS, Network and Applications Trust Layers

- Consistent security must provide security at all layers correspondingly relying on trust credentials at each layer
  - Application – Container - Operating systems (security kernel) + Cloud platform
  - Network/communication – Runtime - Storage
- Two security models: Trusted Computing Base (TCB) for cloud platform and OSI/Internet security cloud based applications
  - Client/server and Service Oriented Architecture vs OS and hypervisor run-time
- Root of trust is based on the security credentials bound to hardware mediated through OS to runtime environment
Cloud Computing Security – Challenges

• Fundamental security challenges and main user concerns in clouds
  – Data security: Where are my data? Are they protected? What control has cloud provider over data security and location?
  – Identity management and access control: Who has access to my personal/ID data?

• Two main tasks in making cloud secure and trustworthy
  – Secure operation of the cloud (provider) infrastructure
  – User controlled access control (security) infrastructure
    • Provide sufficient amount of security controls for competent user

• Cloud security infrastructure should provide a framework for dynamically provisioned cloud security services and infrastructure as a part of the main services
Practical Security Services and Mechanisms Used in Cloud

- **Virtual Private Network (VPN) Virtual Private Cloud (VPC)** for creating virtual cloud infrastructure for each customer
- **Secure Shell (SSH) protocol**
  - Instant SSH keys generation that takes place during a new account creation
    - Private/secret key must be copied/downloaded by user and securely store; CSP keeps only public linked to account, CSP doesn’t store user’s private key
- **HTTPS and TLS/SSL protocols** for secure web access and as a general secure transport protocols
- **Public Key Infrastructure (PKI)** that provides a basis for secure communication protocols (HTTPS, TLS/SSL, SSH) and PKI based credentials for entity/user identification and authentication
- **Access control** that includes Authentication and Authorisation and supported by Identity Management
  - Single Sign On (SSO), federation, delegation, access control policy
- **Federated Access Control and Federated Identity Management**
- **Multi-tenant environment mechanisms** for virtual customer separation
  - Virtualisation, storage and database partitioning, namespace separation, customer session identification and context management, authentication and authorization
- **Key escrow** to ensure restoration of encrypted data in case key held by data owner is lost
Security is declared as one of critical importance to AWS cloud that is targeted to protect customer information and data from integrity compromise, leakage, accidental or deliberate theft, and deletion.

- The AWS infrastructure is designed with the high availability and sufficient redundancy to ensure reliable services operation.
AWS Security – Shared Responsibility Model

AWS implements the *Shared Responsibility Model* that splits responsibility for the security of different layers and components between AWS as a provider and a customer or tenant.

**AWS as a cloud provider** ensures the security of the cloud infrastructure and cloud platform services:
- Facilities
- Physical security of datacentre
- Network infrastructure
- Virtualisation platform and infrastructure

While **the customer** is responsible for security of the following components:
- Amazon Machine instances, OS, and applications
  - Note, the customer is responsible for security update and patching of the guest OS and installed applications
- Data in transit, data at rest, and data stores
- Credentials, policies and configurations
- Comply with the Acceptable Use Policy (AUP), ensure correct use of the cloud platform
Example: Security responsibility sharing in AWS IaaS infrastructure services

- For other cloud service models PaaS and SaaS the responsibility of AWS goes up to OS, network and firewall for PaaS, and also includes the application platform and container for SaaS.
  - However, the responsibility for data remains with the customer.

Recommended **security best practices** at each layer

- Protect your Amazon account
- Control internal access to AWS resources
- Limit external access to your cloud
- Protect data in transit and at rest
- Secure data assets
- Secure your compute assets (OS, instances, App)
- Backup for easy recover
- Keep track of your cloud resources (using monitoring service)

**Security methods** for customer cloud infrastructure

- Virtual Private Cloud (VPC) to create a secure environment for your cloud services in AWS
- Security zoning and network segmentation based on security groups, Network Access Control Lists, host based firewalls
- Network security and secure access for users and applications
- Threats protection layer in traffic flow to ensure protection against Denial of Service (DoS) attacks

Security in Amazon EC2 and S3

Amazon Elastic Compute Cloud (Amazon EC2) Security implements

• Multiple levels of security including Guest Operating System, Firewall, API to manage VM instances
• Hypervisor that is a customised version of the Xen hypervisor allows running processes in four privilege modes: host OS is executed Ring 0; guest OS runs in Ring 1, applications run in Ring 3.
• Instances isolation is also provided by hypervisor that forwards all communication for instances via virtual firewall that resides in the hypervisor layer.
  – Such approach ensures that communication paths of instances never intersect.
• Instant SSH keys generation for individual users and groups

Amazon Simple Storage Service (Amazon S3) Security

• S3 storage is accessed via SSL protocol
• Data security in rest is provided by encryption and multi-layer physical security
• AWS adopts a secure and reliable technique for storage device decommissioning
AWS Identity and Access Management (IAM)

AWS IAM provides functionality to securely control access to AWS services and resources for individual users and groups by defining individual and group permissions and policies.

- Manage **IAM users and their access**: Create AIM users and assign them individual security credentials (i.e., access keys, passwords, and multi-factor authentication devices) to provide users access to the AWS Management Console, APIs, services and resources.
  - Create multiple users and groups for the same AWS account/customer

- Manage **IAM roles and their permissions**: Create roles and manage permissions to control which operations can be performed by the entity

- Manage **federated users and their permissions**: Enable identity federation to allow existing identities (e.g. users) in customer enterprise to access the AWS services and resources, without the need to create an IAM user for each identity.

- Enable **Multi-Factor Authentication (MFA)** that augments username and password credentials
  - MFA requires users to prove physical possession of a hardware or virtual MFA device by providing a valid MFA code.
Example: Multi-layer Security in AWS

Security Layers are defined similar to 3 Tier Model:
- Presentation/Front end
- Applications
- Data/Database

#Permit HTTP(S) access to Web Layer
From entire Internet
ec2auth Web –p 80, 443 –s 0.0.0.0/0

#Permit Web Layer access to App Layer
ec2auth App –p 8000 –o Web

#Permit App Layer access to DB
ec2auth DB –p 3209 –o App

#Permit admin access SSH to all three layers
# First allow connection from office to Web
tier, and from there to the other layers
ec2auth Web –p 22 -s <e.g. office network>
ec2auth App –p 22
ec2auth DB –p 22
Microsoft Azure Cloud Security

Microsoft Azure Cloud is built with the security in mind
- Microsoft has long term experience in developing security applications and running large scale service (e.g., Hotmail, MSDN, MSN, Windows Live).
- Azure cloud design claims to follow the widely recognised Microsoft Secure Software Development Lifecycle (SSDL)

SSDL = Security and Privacy by Design

Three components of the cloud environment security
- Cloud infrastructure security
  - Datacenter security, trustworthy design, secure operational procedures
  - Certification and compliance
- Cloud platform security services
  - Serving both platform security and integration with the customer applications
  - Access control, security policies, customer controlled security services
  - Data protection: cloud platform and user controlled
- Customer/tenants applications security
  - Application level access control, security policies
  - Federation with the tenants/users organizational access control
Microsoft Azure Cloud Security Services

Azure Security Services from both the customer’s and provider’s operational perspectives:

• Federated identity and access management based on Microsoft accounts or organizational accounts, enabled by **Azure Active Directory Service (AADS)**
• Use of mutual SSL authentication.
• Component isolation through a layered environment.
• Virtual Machine state maintenance and configuration integrity.
• Storage redundancy to minimize the impact of hardware failures.
• Monitoring, logging, and reporting on administrative actions.

Built-in data protection

• Control access to customer data and applications
  – Identity and access control with Active Directory based federated access control and Identity Management
• Protect data in transit and at rest
  – Data encryption and isolation
  – Data destruction following string industry standards
• Dedicated network connectivity with Azure ExpressRoute
Big Data Security


  – 10 recommendations are provided for each of CSA Top Ten Big Data Security Challenges
CSA Top Ten Big Data Security and Privacy Challenges

Cloud Security Alliance also published their ‘Expanded Top Ten Big Data Security and Privacy Challenges’ document as early as in June 2013.

Top Ten challenges are grouped into four functional groups:

A. Infrastructure security
   TT01. Secure computations in distributed programming frameworks
   TT03. Secure data storage and transactions logs
   TT04. End-point input validation/filtering
   TT05. Real-time security/compliance monitoring

B. Access control and policy
   TT02. Security best practices for non-relational data stores
   TT08. Granular access control and data centric access policies

C. Data Privacy and Confidentiality
   TT06. Scalable and composable privacy-preserving data mining and analytics
   TT07 Cryptographically enforced data centric security

D. Data Management
   TT09. Granular audits
   TT10. Data provenance
The proposed analysis of the security and privacy challenges includes the following sections:
1. Use cases definition
2. Modeling: Formalizing a threat model that covers most of the cyber-attack or data-leakage scenarios
3. Analysis: Finding tractable solutions based on the threat model
4. Implementation: Suggestions for implementing the solutions in existing infrastructures
Use case:
Most industries and government agencies will benefit from real-time security analytics. Common uses include utilizing the technology to answer questions such as, “Who is accessing which data from which resource at what time,” “Are we under attack,” or “Is there a breach of compliance standard C because of action A?”

5.1 Apply big data analytics to detect anomalous connections to cluster
5.2 Mine logging events
5.3 Implement front-end systems
5.4 Consider cloud-level security
5.5 Utilize cluster-level security
5.6 Apply application-level security
5.7 Adhere to laws and regulations
5.8 Reflect on ethical considerations
5.9 Monitor evasion attacks
5.10 Track data-poisoning attacks
CSA Top Ten TT07. Cryptographically Enforced Data-Centric Security

- CSA Big Data Security and Privacy Handbook Recommendations

Use case:
On-demand provisioned and distributed character of the Big Data infrastructure, especially if it is cloud based, make it practically unfeasible to achieve full protection of data at all infrastructure layers and during the whole data lifecycle, unless data remain encrypted all time.

7.1 Construct system to search, filter for encrypted data
7.2 Secure outsourcing of computation using fully homomorphic encryption
7.3 Limit features of homomorphic implementation
7.4 Apply relational encryption to enable comparison of encrypted data
7.5 Reconcile authentication and anonymity
7.6 Implement identity-based encryption
7.7 Utilize attribute-based encryption and access control
7.8 Use oblivious RAM for privacy preservation
7.9 Incorporate privacy-preserving public auditing
7.10 Consider convergent encryption for deduplication
Use case:
Several key security applications require a digital record with details about its creation. Examples include detecting insider trading for financial companies or determining the accuracy of the data source for research investigations. These security assessments are time-sensitive in nature and require fast algorithms to handle the provenance metadata containing this information. In addition, data provenance complements audit logs for compliance requirements, such as PCI or Sarbanes-Oxley.

10.1 Develop infrastructure authentication protocol
10.2 Ensure accurate, periodic status updates
10.3 Verify data integrity
10.4 Ensure consistency between provenance and data
10.5 Implement effective encryption methods
10.6 Use access control
10.7 Satisfy data independent persistence
10.8 Utilize dynamic fine-grained access control
10.9 Implement scalable fine-grained access control
10.10 Establish flexible revocation mechanisms
Part 2. Cloud Compliance

- Compliance standards, Security Controls
- CSA GRC Stack: Governance, Risk Management and Compliance
- PCI DSS Cloud Computing Guidelines
Security and Compliance

• Security and compliance are related and in some cases interchangeable

• Security is commonly defined as a set of technical, physical, and administrative controls in order to ensure normal operation of a system or application
  – Security is often associated with the CIA triad Confidentiality, Integrity, Availability
  – Appropriate level of security requires organizations to take measures and comply to the numerous security controls

• Compliance is a certification or confirmation that the system or an organization meets the requirements of specified standards, established legislation, regulatory guidelines or industry best practices that can be jointly defined as compliance framework
  – A compliance framework can includes business processes and internal controls the organization has in place to adhere to these standards and requirements
  – The framework should also map different requirements to internal controls and processes to eliminate redundancies

• Why it is important for cloud?
  – When moving to cloud, the organization moves from internal security and operational environment/context (that may not be formally defined) to external operational security that will become a part of SLA (or business requirement) with CSP

• Problem with achieving compliance for cloud based applications/solutions
  – Audit requirements are not designed for virtualised distributed environment
  – Lack of visibility in cloud: large CSP such as Amazon and Google are “walled/curtained gardens”
  – Requirements to allow CSP audit may involve Non-Disclosure Agreement (NDA) and risk of provider lock-in
General standards and recommendations

- ISO/IEC 27001:2005 Certification on security infrastructure
  - Industry standard: the risk-based information security management program that follows a plan-do-check-act process
- NIST SP 800-53 Security Controls and ISO/IEC 15408 Evaluation Criteria
- HIPAA/HITECH - The U.S. Health Insurance Portability and Accountability Act (HIPAA) and Health Information Technology for Economic and Clinical Health (HITECH)
  - Act created by the US federal government include provisions to protect patients' private information.
- NIST SP 800-144 Guidelines for Security and Privacy in Cloud Computing
- ENISA Cloud Computing Security Risk Assessment
Industry and Governmental Regulatory Requirements (USA)

- **Service Organisation Control SOC 1 (SSAE 16/ISAE 3402) and SOC 2 and 3 (AT 101)**
  - SOC 2 is a detailed attestation report (often restricted) for service organizations that contains strict standards for security, availability, processing integrity, confidentiality, and privacy.
  - SOC 3 is a general purpose report which summarizes the SOC 2 audit.

- **Sarbanes Oxley Act (SOX)** also known as "Corporate and Auditing Accountability and Responsibility Act" set enhanced standards for all US public company boards, management and public accounting firms.
  - According to SOX act, top management must individually certify the accuracy of financial information.

  - Describes security requirements for the protection of information and information systems in Federal systems.

- **Department of Defense Information Certification Accreditation Process (DIACAP)**

- **Federal Risk and Authorization Management Program (FedRAMP)**
  - As of June 6, 2014, US federal agencies must utilize only cloud providers assessed and authorized through FedRAMP. List of authorized cloud providers is published:
    - Authorisation: AWS East-West US and AWS Governmental Community Cloud, SalesForce, USDA
    - Provisional Authorisation: Akamai, AT&T, IBM, Hewlett-Packard, Lockheed Martin, Microsoft Azure, Oracle
Mapping Compliance and Cloud Infrastructure Components

Main Security Controls
- Identity/Attribute Management
- End Point Security
- Access Control
- Data Security
- Security Context & Policy
- Network Security
- Key & Session Mgmt
- Conf & Log Mgmt
- Platform Security

Application/Infrastructure Layers
- End User/Enterprise IT Infrastructure
- Customer (Inter)Cloud based Applications & Infrastructure
- Public/Private/Hybrid Cloud Infrastructure

Compliance Domains
- Regulations
- IS/IT Management
- Government
- Finance
- Healthcare
- Energy
The main questions that security and compliance auditors would ask you

- Where is our data going to reside?
- Who is going to look after it?
- Who is going to be able to see it?
- Is it going to be the people that manage the infrastructure for us?
- Is it going to be internal and external people?
- And if we use a public cloud how secure is that cloud platform for us?
- Is the cloud going to be segregated from other organisations’ data?
Case study: Certification/Compliance by Amazon AWS Cloud

The AWS cloud infrastructure has been designed and managed in alignment with regulations, standards, and best-practices including:

- ISO/IEC 27001:2005
- SOC 1, SOC2, SOC3
- FIPS 140-2
- CSA
- PCI DSS Level 1
- HIPAA
- ITAR
- DIACAP and FISMA
- FedRAMP (SM)
- MPAA

Amazon Cloud is certified for hosting US Governmental services

http://aws.amazon.com/compliance/
Microsoft services/infrastructure meets the following key certifications, attestations and compliance capabilities

- ISO/IEC 27001:2005 Certification on security infrastructure
- SOC 1 (SSAE 16/ISAE 3402) and SOC 2 and 3 (AT 101)
  - Obtained in 2008 and 2012
- Cloud Security Alliance (CSA) Cloud Controls Matrix
- NIST SP 800-144 Guidelines for Security and Privacy in Cloud Computing
- PCI Data Security Standard Certification level 1
- HIPAA and HITECH
- FISMA Certification and Accreditation – since 2010
- Various state, federal, and international Privacy Laws (95/46/EC, e.g. EU Data Protection Directive, California SB 1386, etc.)

CSA3.0 Security Guidance for Critical Area of Focus in Cloud Computing

The CSA3.0 defines 13 domains of the security concerns for Cloud Computing that are divided into two broad categories that define corresponding security controls.

**Governance domains**
- 1. Governance and Enterprise Risk Management
- 2. Legal Issues: Contracts and Electronic Discovery
- 3. Compliance and Audit
- 4. Information Management and Data Security
- 5. Portability and Interoperability

**Operational Domains**
- 6. Traditional Security, Business Continuity and Disaster Recovery
- 7. Data Center Operations
- 8. Incident Response, Notification and Remediation
- 9. Application Security
- 10. Encryption and Key Management
- 11. Identity and Access Management
- 12. Virtualization
- 13. Security as a Service
The GRC Stack provides a toolkit for enterprises, cloud providers, security solution providers, IT auditors and other stakeholders to assess both private and public clouds against industry established best practices, standards and critical compliance requirements.

https://cloudsecurityalliance.org/research/grc-stack/

- **Cloud Controls Matrix (CCM)** is designed to provide fundamental security principles to guide cloud vendors and to assist prospective cloud customers in assessing the overall security risk of a cloud provider (https://cloudsecurityalliance.org/research/ccm/)
  - The CCM gives detailed understanding of security concepts and principles that are aligned to the Cloud Security Alliance guidance in 13 domains
  - Defined in accordance to industry-accepted security standards, regulations, and controls frameworks such as the HITRUST CSF, ISO 27001/27002, ISACA COBIT, PCI, HIPAA and NIST.

- **Consensus Assessments Initiative Questionnaire (CAIQ)** provides an industry-accepted way to document what security controls exist in IaaS, PaaS, and SaaS offerings, providing security control transparency (https://cloudsecurityalliance.org/research/cai/)
  - Provided in a form of questionnaire in the spreadsheet format, a set of questions a cloud consumer and cloud auditor may wish to ask of a cloud provider.
  - Questions are formulated to answer “yes or no” what should help to tailor each unique cloud customer’s evidentiary requirements.

- **Other initiatives (currently dormant)**
  - CloudAudit to provide a common interface and namespace that would allow cloud computing providers to automate the Audit, Assertion, Assessment, and Assurance (A6) of their cloud environments and provide a foundation for transparency and trust in private and public cloud systems.
  - Cloud Trust Protocol (CTP) to provide mechanism by which cloud service consumers ask for and receive information about the elements of transparency as applied to cloud service providers.
<table>
<thead>
<tr>
<th>Delivering</th>
<th>Stack Pack</th>
<th>Description</th>
</tr>
</thead>
</table>
| Continuous monitoring ... with a purpose                                   | ![CTP](https://cloudsecurityalliance.org/research/grc-stack/) | **Cloud Trust Protocol (CTP)**  
• Common technique and nomenclature to request and receive evidence and affirmation of current cloud service operating circumstances from cloud providers |
| Claims, offers, and the basis for auditing service delivery                | ![Cloud Audit](https://cloudsecurityalliance.org/research/grc-stack/) | **Common interface and namespace to automate the Audit, Assertion, Assessment, and Assurance (A6) of cloud environments**                                                                                     |
| Pre-audit checklists and questionnaires to inventory controls             | ![CAI](https://cloudsecurityalliance.org/research/grc-stack/)     | **Consensus Assessments Initiative (CAI)**  
• Industry-accepted ways to document what security controls exist                                                                                                                                            |
| The recommended foundations for controls                                  | ![CCM](https://cloudsecurityalliance.org/research/grc-stack/)     | **Cloud Control Matrix (CCM)**  
• Fundamental security principles in specifying the overall security needs of a cloud consumers and assessing the overall security risk of a cloud provider |
CSA3.0: Mapping the Cloud Model to the Security Control & Compliance

Cloud Model

Find the Gaps!

Security Control Model

Applications
- SDLC, Binary Analysis, Scanners, WebApp Firewalls, Transactional Sec.

Information
- DLP, CMF, Database Activity Monitoring, Encryption

Management
- GRC, IAM, VA/VM, Patch Management, Configuration Management, Monitoring

Network
- NIDS/NIPS, Firewalls, DPI, Anti-DDoS, QoS, DNSSEC, OAuth

Trusted Computing
- Hardware & Software RoT & API's

Compute & Storage
- Host-based Firewalls, HIDS/HIPS, Integrity & File/log Management, Encryption, Masking

Physical
- Physical Plant Security, CCTV, Guards

Compliance Model

PCI
- Firewalls
- Code Review
- WAF
- Encryption
- Unique User IDs
- Anti-Virus
- Monitoring/IDS/IPS
- Patch/Vulnerability Management
- Physical Access Control
- Two-Factor Authentication...

HIPAA

GLBA

SOX

SP500-292 (CCRA), CSA3.0

ISO/IEC 27002:2013 InfoSec Controls

PCI DSS V3.0 (2013)

What is the Cloud Controls Matrix (CCM)?

- Baseline control framework specifically designed for managing risk in the Cloud Supply Chain:
  - Addressing the inter and intra-organizational challenges of persistent information security by clearly delineating control ownership.
  - Providing an anchor point and common language for balanced measurement of security and compliance postures.
  - Providing the holistic adherence to the vast and ever evolving landscape of global data privacy regulations and security standards.
- Serves as the basis for new industry standards and certifications.

CCM Control Groups:
1. Compliance (CO)
2. Data Governance (DG)
3. Facility Security (FS)
4. Human Resources (HR)
5. Information Security (IS)
6. Legal (LG)
7. Operations Management (OM)
8. Risk Management (RI)
9. Release Management (RM)
10. Resiliency (RS)
11. Security Architecture (SA)

98 security controls in total
CSA Consensus Assessment Initiative

- A cloud supply chain risk management and due diligence questionnaire
- ~ 200 yes/no questions that map directly to the CCM, and thus, in turn, to many industry standards.
- Can be used by both CSPs for self-assessment or by potential customers for the following purposes
  - to identify the presence of security controls and practices for cloud offerings
  - procurement negotiation
  - contract inclusion
  - to quantify SLAs
- For potential customers, the CSA Consensus Assessment Initiative Questionnaire (CAIQ) is intended to be part of an initial assessment followed by further clarifying questions of the provider as it is applicable to their particular needs.
  - v1.1 published in Sept 2011; v3.0.1 is available from 2014
The following are the principles that the working group utilized as guidance when developing the CAIQ:

- The questionnaire is organized using CSA 13 governing & operating domains divided into “control areas” within CSA’s Control Matrix structure.
- Questions are to assist both cloud providers in general principles of cloud security and clients in vetting cloud providers on the security of their offering and company security profile.
- CAIQ not intended to duplicate or replace existing industry security assessments but to contain questions unique or critical to the cloud computing model in each control area.
- Each question should be able to be answered yes or no.
- If a question can’t be answered yes or no then it was separated into two or more questions to allow yes or no answers.
- Questions are intended to foster further detailed questions to provider by client specific to client’s cloud security needs. This was done to limit number of questions to make the assessment feasible and since each client may have unique follow-on questions or may not be concerned with all “follow-on questions.”
CSA STAR: Security, Trust and Assurance Registry

- Public Registry of Cloud Provider self assessments
  - [https://cloudsecurityalliance.org/star/#_registry](https://cloudsecurityalliance.org/star/#_registry)
- Leverages GRC Stack Projects
  - Consensus Assessments Initiative Questionnaire
  - Provider may substitute documented Cloud Controls Matrix compliance
- Voluntary industry action promoting transparency
- Free market competition to provide quality assessments
- Documents the security controls provided by various cloud computing offerings
- Encourage transparency of security practices within cloud providers
- Permanent effort to drive transparency, competition, innovation and self regulation with agility – crowdsourcing cloud security

Open Certification Framework

3. Continuous Monitoring-Based Certification
2. 3rd-Party Assessment-Based Certification
1. Self-Assessment

CSA STAR Assessment and Certification

Cloud Controls Matrix

Consensus Assessments Initiative

BDIC2016 Big Data Security and Compliance 56
CSA STAR Compliance Levels

LEVEL ONE: CSA STAR Self-Assessment
• CSA STAR Self-Assessment is a free offering that documents the security controls provided by CSPs, thereby helping users assess the security of cloud providers
• Cloud providers either submit a completed The Consensus Assessments Initiative Questionnaire (CAIQ), or to submit a report documenting compliance with Cloud Controls Matrix (CCM).

LEVEL TWO: CSA STAR Attestation
• CSA STAR Attestation is a collaboration between CSA and the AICPA to provide guidelines for CPAs to conduct SOC 2 engagements using criteria from the AICPA (Trust Service Principles, AT 101) and the CSA Cloud Controls Matrix.

LEVEL TWO: CSA STAR Certification
• The CSA STAR Certification is a rigorous third party independent assessment of the security of a cloud service provider.
• The technology-neutral certification leverages the requirements of the ISO/IEC 27001:2005 management system standard together with the CSA Cloud Controls Matrix.

LEVEL THREE: CSA STAR Continuous Monitoring
• Currently under development and scheduled for 2015 release, CSA STAR Continuous Monitoring enables automation of the current security practices of cloud providers.

Listing at https://cloudsecurityalliance.org/star/#star_m
STAR Listing Process

- Provider fills out CAIQ or customizes CCM
- Uploads document at /star repository
- CSA performs basic verification
  - Authorized listing from provider
  - Delete SPAM, “poisoned” listing
  - Basic content accuracy check
- CSA digitally signs and posts at /star
- Does not provide automation, 3rd party assessment, relative/absolute scoring, real-time controls monitoring, etc
- Ultimate assurance is real time GRC (enabled by CloudAudit) complemented by CSA STAR and 3rd party attestation.
Why not certification or 3rd party assessment?

• Complex to do certification right
  – Many uses of cloud, many customer needs
  – Different risk profiles for each

• CSA supporting broad industry consortia and standards bodies
  – ISO, ITU-T
  – Common Assurance Maturity Model (CAMM – 3rd Party assessment)
  – GRC Stack aligns with common requirements (e.g. PCI/DSS, HIPAA, FedRAMP, 27001, CoBIT, etc)

• Self assessment & transparency complements all
  – STAR could be part of SSAE 16 SOC II report (SAS 70 replacement)
Payment Card Industry Data Security Standard (PCI DSS) and Compliance
**PCI DSS Principles and Requirements (Version 3.0, 2013)**

<table>
<thead>
<tr>
<th>Principle</th>
<th>Requirements</th>
</tr>
</thead>
</table>
| **Build and Maintain a Secure Network**       | Requirement 1: Install and maintain a firewall configuration to protect cardholder data  
Requirement 2: Do not use vendor-supplied defaults for system passwords and other security parameters                                    |
| **Protect Cardholder Data**                   | Requirement 3: Protect stored cardholder data  
Requirement 4: Encrypt transmission of cardholder data across open, public networks                                                   |
| **Maintain a Vulnerability Management Program** | Requirement 5: Use and regularly update anti-virus software  
Requirement 6: Develop and maintain secure systems and applications                                                                            |
| **Implement Strong Access Control Measures**  | Requirement 7: Restrict access to cardholder data by business need-to-know  
Requirement 8: Assign a unique ID to each person with computer access  
Requirement 9: Restrict physical access to cardholder data                                                                                   |
| **Regularly Monitor and Test Networks**       | Requirement 10: Track and monitor all access to network resources and cardholder data  
Requirement 11: Regularly test security systems and processes                                                                                |
| **Maintain an Information Security Policy**   | Requirement 12: Maintain a policy that addresses information security                                                                         |

- PCI DSS is a multifaceted security standard that includes requirements for security management, policies, procedures, network architecture, software design and other critical protective measures.
- Other PCI standards include: PIN Entry Device (PED), DSS Self-Assessment, Payment Application DSS, and Qualified Security Assessment (QSA) Program

- Cloud Overview – provides explanation of common deployment and service models for cloud environments, including how implementations may vary within the different types.
- Cloud Provider/Cloud Customer Relationships – outlines different roles and responsibilities across the different cloud models and guidance on how to determine and document these responsibilities.
- PCI DSS Considerations – provides guidance and examples to help determine responsibilities for individual PCI DSS requirements, and includes segmentation and scoping considerations.
- PCI DSS Compliance Challenges - describes some of the challenges associated with validating PCI DSS compliance in a cloud environment.
- Additional Security Considerations – explores a number of business and technical security considerations for the use of cloud technologies.
Audience: Involved parties/stakeholders

Cloud customers, cloud service providers and regulators must work together to determine the solutions that best meet the needs of stakeholders:

- **Merchants**: Enterprise and public sector cloud customers must be able to achieve their compliance obligations while using cloud services
- **Individuals** using the cloud have an expectation that their personal information will be protected and used appropriately
- **Cloud service providers** must have clear mechanisms to evaluate and communicate capabilities
- **Assessors**: Regulators and industry governing bodies need to have confidence that their requirements are met and verified
Example: Controls assignment for different cloud service models (PCI DSS)

CSP facility/infrastructure must be certified/compliant with PCI DSS and relevant standards
- Provided mapping between PCI DSS and other standards

Note: Data is always responsibility of the customer/client

Source: PCI DSS Cloud Computing Guidelines, February 2013
(https://www.pcisecuritystandards.org/pdfs/PCI_DSS_v2_Cloud_Guidelines.pdf)
### Example: PCI DSS Responsibilities sharing

<table>
<thead>
<tr>
<th>PCI DSS Requirement</th>
<th>Example responsibility assignment for management of controls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1:</strong> Install and maintain a firewall configuration to protect cardholder data</td>
<td>IaaS: Both, PaaS: Both, SaaS: CSP</td>
</tr>
<tr>
<td><strong>2:</strong> Do not use vendor-supplied defaults for system passwords and other security parameters</td>
<td>IaaS: Both, PaaS: Both, SaaS: CSP</td>
</tr>
<tr>
<td><strong>3:</strong> Protect stored cardholder data</td>
<td>IaaS: Both, PaaS: Both, SaaS: CSP</td>
</tr>
<tr>
<td><strong>4:</strong> Encrypt transmission of cardholder data across open, public networks</td>
<td>IaaS: Client, PaaS: Both, SaaS: CSP</td>
</tr>
<tr>
<td><strong>5:</strong> Use and regularly update anti-virus software or programs</td>
<td>IaaS: Client, PaaS: Both, SaaS: CSP</td>
</tr>
<tr>
<td><strong>6:</strong> Develop and maintain secure systems and applications</td>
<td>IaaS: Both, PaaS: Both, SaaS: Both</td>
</tr>
<tr>
<td><strong>7:</strong> Restrict access to cardholder data by business need to know</td>
<td>IaaS: Both, PaaS: Both, SaaS: Both</td>
</tr>
<tr>
<td><strong>8:</strong> Assign a unique ID to each person with computer access</td>
<td>IaaS: Both, PaaS: Both, SaaS: Both</td>
</tr>
<tr>
<td><strong>9:</strong> Restrict physical access to cardholder data</td>
<td>IaaS: CSP, PaaS: CSP, SaaS: CSP</td>
</tr>
<tr>
<td><strong>10:</strong> Track and monitor all access to network resources and cardholder data</td>
<td>IaaS: Both, PaaS: Both, SaaS: CSP</td>
</tr>
<tr>
<td><strong>11:</strong> Regularly test security systems and processes</td>
<td>IaaS: Both, PaaS: Both, SaaS: CSP</td>
</tr>
<tr>
<td><strong>12:</strong> Maintain a policy that addresses information security for all personnel</td>
<td>IaaS: Both, PaaS: Both, SaaS: Both</td>
</tr>
</tbody>
</table>

**PCI DSS Appendix A: Additional PCI DSS Requirements for Shared Hosting Providers**

CSP, IaaS, PaaS, SaaS
Segmentation considerations

Segmentation on a cloud based infrastructure must ensure similar level of segmentation as with physical infrastructure

- Mechanisms to ensure appropriate segmentation may be applied at the network, operating system and application layers
- Cardholders data isolation must be ensured
- (Multi-tenant) Cloud environments must be isolated from each other

Examples of correctly segmented cloud environments:

- Traditional Application Service Provider (ASP) model where physically separate servers are provided for each client’s environment
- Virtualized servers that are individually dedicated to a particular client, including any virtualized disks such as SAN, NAS or virtual database servers
- Environments where clients run their applications in separate logical partitions using separate database management system images and do not share disk storage or other resources

Examples of not correct segmentation

- Environments are separated only by access control at the application or operating system level
- Data are stored at the same of the database
Segmentation technologies

- Physical firewalls and network segmentation at the infrastructure level
- Firewalls at the hypervisor and VM level
- VLAN tagging, in addition to firewalls
- Intrusion prevention system
- Data loss prevention tools at the hypervisor and/or VM level
- Controls to prevent out-of-band communications
- Isolation of shared processes and resources from client environments
- Segmented data storage for each client
- Strong, two-factor authentication
- Separation of duties and administrative oversight
- Continuous logging and monitoring of perimeter traffic
What does PCI DSS Compliance means?

- Cloud service validation for certain PCI DSS requirements does not automatically transfer to the client environments within the cloud service
  - For example, a CSP’s validation may have included use of up-to-date anti-virus software on the CSP’s systems; however, this validation might not extend to the individual client OS or VMs (such as in an IaaS service).
- Client’s PCI DSS compliance does not result in any claim of compliance for the CSP, even if the client’s validation included elements of the service managed by the CSP.

In general, one party’s compliance doesn’t solve the overall compliance
a) If a CSP is compliant, this does not mean that their clients are.
b) If a CSP’s clients are compliant, this does not mean that the CSP is.
c) If a CSP and the client are compliant, this does not mean that any other clients are

SLA and compliance
- SLA and other written agreements should clearly delineate responsibility between parties
- PCI DSS compliance validation and testing (with associated controls and permissions) should be clearly detailed in SLA
Each PCI DSS SAQ consists of the following components:

1. Questions correlating to the PCI DSS requirements, as appropriate for different environments: See “Selecting the SAQ and Attestation that Best Apply to Your Organization” in this document. This section also includes a column for “Expected Testing” which is based on the testing procedures in PCI DSS.

2. Attestation of Compliance: The Attestation includes your declaration of eligibility for completing the applicable SAQ and the subsequent results of a PCI DSS self-assessment.
SAQ types

- **SAQ A** – Card-not-present Merchants, All Cardholder Data Functions Fully Outsourced
- **SAQ A-EP** – Partially Outsourced E-Commerce Merchants Using a Third-Party Website for Payment Processing
- **SAQ B** – Merchants with Only Imprint Machines or Only Standalone, Dial-Out Terminals. No Electronic Cardholder Data Storage
- **SAQ B-IP** – Merchants with Standalone, IP-Connected PTS Point-of-Interaction (POI) terminals, No Electronic Cardholder Data Storage
- **SAQ C-VT** – Merchants with Web-Based Virtual Terminals, No Electronic Cardholder Data Storage
- **SAQ C** – Merchants with Payment Application Systems Connected to the Internet, No Electronic Cardholder Data Storage
- **SAQ P2PE** – Merchants using Only Hardware Payment Terminals in a PCI SSC-listed P2PE Solution, No Electronic Cardholder Data Storage
- **SAQ D for Service Providers** – SAQ-Eligible Service Providers
- **SAQ D for Merchants** – All Other SAQ-Eligible Merchants
Part 4. Hands On Exercise/Lab

Cloud Compliance Assessment tools

• CSA CAIQ and PCI DSS Dashboard
  – CSA Cloud Controls Matrix (CCM) v3.0.1
    https://cloudsecurityalliance.org/research/ccm/
    https://downloads.cloudsecurityalliance.org/initiatives/ccm/ccm-v3.0.1.zip
  – CSA Consensus Assessment Initiative Questionnaire (CAIQ) v3.0.1
    https://cloudsecurityalliance.org/research/cai/
    https://cloudsecurityalliance.org/download/consensus-assessments-initiative-questionnaire-v3-0-1/

• PCI DSS Self Assessment Questionnaire (SAQ)
  – Questionnaire is designed for different categories of user: from full outsourcing cards operations to card payment service
Part 3. Privacy protection and regulations

- Privacy related standards
- EU General Data Protection Regulation (GDPR)
Security and compliance standards are also covering data protection and privacy. Specifically focused documents:

- The White House report ‘Big Data: Seizing Opportunities, preserving values’ (May 2014)
- Consumer Data Privacy In A Networked World: A Framework For Protecting Privacy And Promoting Innovation In The Global Digital Economy (February 2012)
- HIPAA/HITECH - The U.S. Health Insurance Portability and Accountability Act (HIPAA) and HITECH (Health Information Technology for Economic and Clinical Health)
  - Act created by the US federal government include provisions to protect patients' private information.

- Protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), Brussels, 25.1.2012, COM(2012)
- Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data
- EU – U.S. Privacy Shield, 12 July 2016: Protection of Transatlantic data flows
  - Repeals former U.S.-EU Safe Harbor existed until 2016
White House report “Big Data: Seizing Opportunities, preserving values” (2014)

• The White House report ‘Big Data: Seizing Opportunities, preserving values’ published in May 2014.
  – The report is the result of the 90-day study commissioned by the President of the United States to examine how big data will transform the way people live and work and how big data will alter the relationships between government, citizens, businesses and consumers.

• Data security and privacy challenges in Cloud Computing and big data have been a focus of numerous study groups initiated by different governmental bodies that produced several valuable reports.
  – Wide implementation of Cloud Computing provided a basis for developing big data technologies and data-centric and data-driven applications that in their own turn facilitate cloud technologies development.

• **The main approach in developing recommendations was to protect privacy while not hindering/restricting development of new technology for the benefit of the whole society.**
  – The report expresses the opinion that despite widely discussed needs for personal control of the collected e-commerce and social data, the practical use of such control is impractical due to the unmanageable volume of information and its variety. Instead, the advertisement companies and other organisational users of the personally
A number of recent documents developed by President Obama’s administration provide a foundation for the Right for Privacy in our Information society. A key document in this respect is the ‘Consumer Data Privacy In A Networked World: A Framework For Protecting Privacy And Promoting Innovation In The Global Digital Economy’ published in February 2012, which states the following rights:

- **Individual Control**: Consumers have a right to exercise control over what personal data organisations collect from them and how they use it.
- **Transparency**: Consumers have a right to easily understandable information about privacy and security practices.
- **Respect for Context**: Consumers have a right to expect that organisations will collect, use, and disclose personal data in ways that are consistent with the context in which consumers provide the data.
- **Security**: Consumers have a right to secure and responsible handling of personal data.
- **Access and Accuracy**: Consumers have a right to access and correct personal data in usable formats, in a manner that is appropriate to the sensitivity of the data and the risk of adverse consequences to consumers if the data are inaccurate.
- **Focused Collection**: Consumers have a right to reasonable limits on the personal data that companies collect and retain.
- **Accountability**: Consumers have a right to have personal data handled by companies with appropriate measures in place to assure they adhere to the Consumer Privacy Bill of Rights.

The theory behind the Consumer Privacy Bill of Rights and other documents is that the combination of broad baseline principles and specific codes of conduct can protect consumers while supporting innovation.
Special Eurobarometer 431 - Data protection, June 2015

- A large majority of people (71%) still say that providing personal information is an increasing part of modern life and accept that there is no alternative other than to provide it if they want to obtain products or services.
- Over half of Europeans who use the Internet use an online social network at least once a week. This proportion is similar for using messaging or chat sites.
- A large majority of Europeans (69%) would like to give their explicit approval before the collection and processing of their personal data.
- More than six out of ten respondents say that they do not trust landline or mobile phone companies and internet service providers (62%) or online businesses (63%).
- 67% find it important to be able to transfer personal data to a new online service provider (‘data portability’).

Trust in internet services:

- 81% of Europeans feel that they do not have complete control over their personal data online.
- Only 24% of Europeans have trust in online businesses such as search engines, social networking sites and e-mail services.
EU documents outlining EU regulatory basis for data protection in cloud

• Former Framework: EU data protection law is based on Directive 95/46/EC (the “Directive”), which was introduced in 1995.
  – Significant technology advances, and fundamental changes to the ways in which individuals and organisations communicate and share information.
  – Many EU Member States have taken divergent approaches to implementing the Directive, creating compliance difficulties for many businesses

• *Proposal* for Regulation on on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), Brussels, 25.1.2012, COM(2012) (119 pages)

• *NEW!* Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Published 4 May 2016 (88 pages)
  – To enter into force in all Member States from 25 May 2018, “without delay, any subsequent amendment affecting them.”
GDPR main principles and key changes

- One simple technologically neutral and future-proof set of rules across the EU
  - Help building trust in the online environment (that is global, distributed, opaque)

- Everyone has the right to the protection of personal data
  - Under EU law, personal data can only be gathered legally under strict conditions, for a legitimate purpose.
  - Furthermore, persons or organisations which collect and manage your personal information must protect it from misuse and must respect certain rights of the data owners which are guaranteed by EU law.

- Privacy impact and requirements
  - Right to be forgotten (RTBF) – complex issue for global cloud infrastructures and all information collected on the web and mobile applications

- Data portability: right to obtain data and relocate to new provider/location
- Recommendation to establish a position responsible for compliance
- Privacy by-design principle for services and infrastructure
EU – U.S. Privacy Shield

- EU-U.S. Privacy Shield: stronger protection for transatlantic data flows
  - Adopted 12 July 2016. Repeals former Safe Harbor Framework

- The new framework protects the fundamental rights of anyone in the EU whose personal data is transferred to the United States as well as bringing legal clarity for businesses relying on transatlantic data transfers.

- The new arrangement includes:
  - strong data protection obligations on companies receiving personal data from the EU
  - safeguards on U.S. government access to data;
  - effective protection and redress for individuals;
  - annual joint review to monitor the implementation.

- What will it mean in practice for American companies
  - Self-certify annually that they meet the requirements.
  - Display privacy policy on their website.
  - Reply promptly to any complaints.
  - If handling human resources data: Cooperate and comply with European Data Protection Authorities.
Impact of GDPR on Businesses (1)


- **Greater harmonisation**: The Regulation introduces a single-legal framework that applies across all EU Member States without the need for national implementation. Businesses will face a *more consistent set of data protection obligations* from one EU Member State to the next.

- **The risk-based approach to compliance**: The Regulation acknowledges a risk-based approach to compliance. *Low-risk processing activities face a reduced compliance burden.* On the other hand, documented *data protection impact assessments* will be required for high-risk processing activities.

- The ‘One-Stop Shop’ for business established in more than one EU Member State.

- **Pseudonymisation**: The concept of *pseudonymised data* (i.e., key-coded or enhanced data) with necessary protection of the mapping keys.

- **Binding Corporate Rules (“BCRs”)**: The Regulation formally recognises BCRs. They will still require SA approval, but the approval process should become *less onerous than the current system*. BCRs are available to both *controllers and processors*. 
Impact of GDPR on Businesses (2)

- **Increased enforcement powers:** The Regulation will significantly increase the maximum fine to €20 million, or 4% of annual worldwide turnover, whichever is greater. Currently, fines under EU Member State law vary, and are comparatively low (e.g., the UK maximum fine is £500,000).

- **Consent, as a legal basis for processing, will be harder to obtain:** Individuals’ consent must be freely given, specific, informed and unambiguous, not bundled and not ‘mere acquiescence’

- **Data protection by design and by default:** Businesses will also be required to perform data protection impact assessments to identify privacy risks in new products.

- **Expanded territorial scope:** Many non-EU businesses that were not required to comply with the Directive will be Strict data breach notification rules: The Regulation will require businesses to notify the SA of data breaches within 72 hours. If potentially serious harm, individuals will have to be notified without undue delay.

- **The ‘right to be forgotten’:** Individuals will have the right to request that businesses delete their personal data in certain circumstances (e.g., the data is no longer necessary for purposes for which it was collected).

- **The right to object to ‘profiling’:** Under the Regulation, individuals will have the right to object to profiling on grounds relating to their particular situation.
• **Expanded territorial scope:** Many non-EU businesses that were not required to comply with the Directive will be required to comply with the Regulation. Non-EU businesses will be subject to the Regulation if they: (i) offer goods or services to EU residents; or (ii) monitor the behaviour of EU residents.

• **Data Protection Compliance Programmes — Internal processing records and Data Protection Officer (DPO):** must be implemented.

• **New obligations of processors:** The Regulation introduces direct compliance obligations for processors. The processors may be liable to pay fines of up to €20 million, or 4% of annual worldwide turnover, whichever is greater.

• **The right to Data Portability:** Individuals will have the right to obtain a copy of their personal data from the controller in a commonly-used format and have it transferred to another controller.
Summary and take away

- Cloud compliance provides a basis for wider cloud services adoption and inter-cloud integration.
- Compliance is supported by numerous standards, legislation, regulatory guidelines and industry best practices that jointly define a compliance framework
  - Knowing major cloud compliance standards is necessary for correct cloud services design, deployment and operation
- PCI DSS Cloud Computing Guidelines provides invaluable source best practices information and practical recommendations for all categories of cloud specialists and practitioners
- Data Protection and Privacy in cloud is regulated by numerous group of standards and regulatory documents
  - European General Data Protection Regulation (GDPR) provides common framework for all EU Member States
  - EU-U.S. Privacy Shield is a new framework for cross-atlantic cooperation
General standards and regulatory requirements related to security and privacy

General Regulatory Requirements for Cloud Compliance


The major industry and government related documents:

- Service Organisation Control SOC 1 (SSAE 16/ISAE 3402) and SOC 2 and 3 (AT 101) (http://www.aicpa.org/interestareass/frc/assuranceadvisoryservices/pages/serviceorganization'smanagement.aspx), including
- Sarbanes Oxley Act (SOX, https://www.sec.gov/about/laws/soa2002.pdf) also known as "Corporate and Auditing Accountability and Responsibility Act"
- HIPAA/HITECH (The U.S. Health Insurance Portability and Accountability Act (HIPAA) and HITECH (Health Information Technology for Economic and Clinical Health) http://www.hhs.gov/ocr/privacy/hipaa/administrative/statute/hipaastatute.pdf)
- Department of Defense Information Certification Accreditation Process (DIACAP) (http://www.prim.osd.mil/Documents/DIACAP_Slick_Sheet.pdf)
• Cloud Security Alliance [https://cloudsecurityalliance.org/](https://cloudsecurityalliance.org/)
    [https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Expanded_Top_Ten_Big_Data_Security_and_Privacy_Challenges.pdf](https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Expanded_Top_Ten_Big_Data_Security_and_Privacy_Challenges.pdf)

• European Union Agency for Network and Information Security
  – ENISA Threat Landscape 2013, Overview of current and emerging cyber-threats, 11 December 2013

• U.S.-EU Safe Harbor List, U.S. Government [online]
  [http://www.export.gov/safeharbor/eu/eg_main_018365.asp](http://www.export.gov/safeharbor/eu/eg_main_018365.asp)
  – U.S.-EU Safe Harbor Framework Documents [online]
    [http://www.export.gov/safeharbor/eu/eg_main_018493.asp](http://www.export.gov/safeharbor/eu/eg_main_018493.asp)
European Union Cloud Computing and Data Protection Directives and Regulations

- Protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), Brussels, 25.1.2012, COM(2012)
- Cloud Service Level Agreement Standardisation Guidelines, Brussels, 24 June 2014 [online]
- Cloud Computing Certification Schemes List – CCSL
  [Link](https://resilience.enisa.europa.eu/cloud-computing-certification)
- New EU GDPL – 4 May 2016 – 4 May 2018
The following documents are the main documents cited in the White House Big Data Privacy report that are related to privacy protection in our networked and IT driven world that experiences Big Data technologies emergence.

In order to bridge differences in approach and provide a streamlined means for U.S. organizations to comply with the EU Directive, the U.S. Department of Commerce in consultation with the European Commission developed a so-called "Safe Harbor" framework that defines key principles and provides recommendations for companies to protect personal data and personal information

- "Personal data" and "personal information" are data about an identified or identifiable individual that are recorded in any form.

7 Safe Harbor privacy principles

- NOTICE: Inform individuals about the purposes for data collection
- CHOICE: Individuals can choose (opt out) whether their personal information is (a) to be disclosed to a third party, or (b) to be used for other purpose.
- ONWARD TRANSFER: To disclose information to a third party, apply the Notice and Choice Principles.
- SECURITY: Personal information must be protected from loss, misuse and unauthorized access, disclosure, alteration and destruction.
- DATA INTEGRITY: Personal information must be relevant for the intended use.
- ACCESS: Individuals must have access to personal information that an organization holds and be able to correct, amend, or delete.
- ENFORCEMENT: Mechanisms for assuring compliance with the Principles, and consequences.
The White House report ‘Big Data: Seizing Opportunities, preserving values’ (May 2014)

- The report contains a comprehensive overview of the data protection and privacy regulations in the United States in their historical development and high-level recommendations for how to address privacy issues in big data applications.
- The document also refers to the historically first article by Samuel Warren and Louis Brandeis in 1890 ‘The Right to Privacy’ related to personal privacy that noted:
  - ‘[r]ecent inventions and business methods call attention to the next step which must be taken for the protection of the person, and for securing to the individual . . . the right ‘to be let alone’ . . . numerous mechanical devices threaten to make good the prediction that ‘what is whispered in the closet shall be proclaimed from the house-tops.’”
  - The report analysed ‘harmful consequences that might result from automated personal data systems’ and recommended certain safeguards when using such information in computer systems (Report, 1973).
  - Those safeguards, commonly known today as the Fair Information Practice Principles (FIPPs), form the basis of modern data protection regulation (Executive Office of the President, 2014).