Cloud Privacy By Design

Open by Default, Private By Design

Introduction - Building the European G-Cloud

The objective of this paper is to describe ‘Cloud Privacy By Design’ (C-PbD), as part of developing the theme as one chapter of a larger white paper: Building the European G-Cloud, with the view it can meet the Cloud privacy needs of the EU public sector.

It will also be published as a standalone document, to make available C-PbD as a general best practice for any other industry to use the same way.

Innovation and Adoption

We can see the EU activities defining this field in two main ways: i) their own planned adoption of Cloud services they will expect to see from local Cloud providers, and ii) the associated research topics being funded by their Horizon 2020 program.

Naturally the best academic and commercialization results for ii) will come from a linked dynamic with i), and the purpose of this paper is to establish that link and use it as a framework for developing C-PbD.

This program is effectively summarized in their H2020 ICT call document, ICT 8, 2015 - Boosting public sector productivity and innovation through cloud computing services:

"Cloud computing has the potential to reduce the IT expenditure of the public sector across Europe, while at the same time improving the scope for flexible high-quality new services. However, the adoption of cloud computing services by the public sector is inhibited by many barriers related to procurement, trustworthiness, technical standards and legal terms of reference, risk of lock-in etc. The overall challenge is to overcome these barriers in order to boost the public sector's productivity by stimulating the preparedness for wide adoption of
Executive Overview

To provide a comprehensive framework the C-PbD program will be defined by:

- Privacy By Design
- Cloud Encryption and Key Management
- Identity and Access Control
- Archiving
- Audit and Certification

Privacy By Design

Privacy By Design (PbD) provides the overarching framework for approaching the use of Cloud services,

ie. although concerns are often described in the form of security issues (hackers et al), it is ultimately user-defined privacy that is the principle objective, with the perspective of ‘Open By Default, Private By Design’ addressing the factor that sometimes too much of a blunt security approach is applied with lack or difficulty of information access for users becoming an issue.

In short the goal is that users control of their own information and also its’ security are both total, and with PbD originating from the public sector it is also well positioned to address the other fundamental goal for C-PbD: Compliance.

Industries like Government and Finance require complete documentation and auditability to be able to safely use Cloud services and also supporting legislation. The goals of this program will be to understand and document where these supports may need further development.

Cloud Encryption and Key Management

Encryption of data is generally considered to be the foundation required for secure use of the Cloud. Relevant standards include FIPS-140.

Horizon 2020

The Horizon 2020 program has an extensive focus on encryption.

From page 68, ICT Call 32, it describes the EUs requirements for encryption technologies, through ‘Trustworthy ICT’:
“Research projects have to address the key challenges to guarantee the security for the lifespan of the application it supports, to stay ahead of the evolution of the ICT environment and keep pace with the performance increase of ICT technology. The challenges to be addressed include:

- Resource efficient and highly secure technology for hardware based real-time cryptography;
- Resource efficient, real-time, highly secure fully homomorphic cryptography;
- Distributed cryptography including functional cryptography;
- Cryptographic tools for securely binding applications to software, firmware and hardware environments, with or without the possibility to adapt the cryptographic primitives which are used;
- Post-quantum cryptography for long term security;
- Quantum key distribution (QKD) systems and networks for long-term security by design, as well as networks supporting information theoretic cryptographic primitives, including but not restricted to the integration into existing optical networks (e.g. trusted nodes and/or using multiplexing), addressing:
  i. low cost components for short-distance, low-bit-rate quantum key-distribution;
  ii. high-bit rate QKD systems that are tolerant to noise and loss. Projects have to demonstrate a net increase in performance, or reduction in energy or power consumption, compared to state-of-the-art approaches and have to validate the proposed technology in realistic application scenarios, taking into account the current trends in ICT like cloud, mobile, IoT, etc. Activities may include methods for provable security against physical attacks, as well as research toward security certification.”

There is also a call for Co-ordination support:

"b. Activities supporting the Cryptography Community

To complement the research activities in cryptography support and coordination actions should address the following aspects:

- ensure a durable integration and structuring of the European cryptography community, involving academia, industry, law enforcement and defence agencies.
- strengthen European excellence in this domain.
- provide technology watch, joint research agendas and foresight studies.
- identify technology gaps, market and implementation opportunities.
- provide technical expertise to the cybersecurity and privacy communities.
- contribute to the development of European standards, including for the public sector.
- solve training needs and skill shortage of academia and industry.
- evaluation and verification of cryptographic protocols and algorithms.
- organize open competitions with security and implementation benchmarking.
- dissemination and outreach, strengthening the link with institutional stakeholders.

Key Management

A critical aspect is where and how encryption keys are stored. Relevant open standards include OASIS KMIP.

Vendor Mapping

These requirements provide a framework for identifying where and how vendors might play different roles in implementing the best practices.

For example vendors like Perspecsys provide on-premise devices that encrypts information before uploading it to Cloud services like Salesforce.com. The Ontario Privacy Commissioner describes:

“PerspecSys’ cloud solution advances the essential principles of Privacy by Design -privacy by default, end-to-end security, and strong accountability- by enabling enterprises to automatically encrypt (decrypt) data locally and on-the-fly when using third-party cloud service applications, with no loss of functionality. Maintaining control over encryption and decryption processes is an excellent approach to safeguarding personal data, regardless of where the data resides. - Ann Cavoukian, Ph.D. Information and Privacy Commissioner of Ontario, Canada”

Other vendors further describe where their technologies can be applied:

- PCI Compliance: HyCloud
- CipherCloud provides a discovery tool
- Porticor offers a solution for another specific vendor VMware
- Disk encryption – Vendors like EMC RSA cater for encrypting all data at the disk level.

Identity Access Control
Federated Identity

Another building block of C-PbD is ‘Federated Identity’.

This enables access to web resources through sites sharing the users credentials, a field that is now quickly maturing with governments like Canada, UK, USA, New Zealand and many more standardizing on protocols like SAML and OAuth to begin building an Identity-centric system for sharing data between agencies.

As explained in [this research](#), and this [Open Identity Exchange presentation](#), Kantara and UMA builds on core Identity building blocks like OpenID Connect to facilitate Identity-enabled secure data sharing between remote systems, facilitating the Personal Identity Ecosystem.

This presentation explains these core mechanics, and this [one maps it to Government Privacy policies](#).

This includes upgrading consulting practices like Privacy Impact Assessments (PIAs) to their Cloud equivalents, the Federated PIA, and modelling the same functions of Cloud Identity Service Providers, Access Control Service Providers and Data Service Providers.

Personal Cloud Ecosystems

This ecosystem is explained in this report from the World Economic Forum – [Personal Data : The Emergence of a New Asset Class](#) (40-page PDF), and where this overlaps with Cloud Computing it’s called ‘Personal Clouds’.

This refers to the ‘Personal Data Stores’ of information we store about ourselves with different online providers, from email through to social networks and photo stores. As well as the Cloud storing the files and data, there will also be protocols and standards for linking these ‘fragments’ together into one holistic digital persona.

This is introduced in this video of the "[Personal Data Locker](#)".

How this new approach might be achieved is introduced in [this white paper](#) from Ontario Privacy Commissioner Ann Cavoukian.

Co-authored by the inventor of the [OASIS XDI protocol](#) this explains the role these Identity standards will play in enabling a new approach based on ‘Personal Clouds’.
On page 13 the author describes ‘Data-by-Reference’, one of a number of architectural shifts of this trend, calmly laid out as part of a best practices guide but representing a major shift in enterprise IT design.

5.4 Cloud Archiving

A critical requirement for the public sector is compliance. Not only does information need to be secured through encryption, but record keeping must also be achieved in line with legislative requirements, prompting the need for the domain of ‘Cloud Archiving’.

This end to end scope is well defined in the CIO Council document Best Practices for Acquiring IT as a Service, a white paper that stipulates a best practice framework for buying services from Cloud Providers, covering a range of topics such as contracting, records management, compliance, legal issues and security, so that agencies can be suitably diligent in procuring their first Cloud projects.

Cloud Storage

This contextualizes the field of Cloud Storage, and therefore a key standards organization is SNIA, who offer relevant presentations including:

- [Cloud Storage](#)
- [Cloud Archiving](#)
- [Re-thinking Archiving](#)
- [Challenges and best practices](#)

Cloud Archiving technical standard are defined in these presentations, which explains that via the ‘XAM’ standard Cloud Archiving builds on Cloud Storage by adding key functionality such as:

- Overwrite protection
- Time or event based retention policy
- Deletion per retention policy
- Archiving metadata
- Full Search capabilities

When records are transferred from one CSP environment to another, the agency will need to be able to ensure the authenticity and completeness of the data they receive. Federal agencies will also need to ensure that all records are deleted from the previous CSP environment once the transition is completed.

Interoperability
Other open standards include those relevant to interoperability, at this records level. I.e. agencies may want to move records between different Cloud providers, not applications or data.

- **CMIS** is an OASIS vendor-driven standard for content management interoperability services.
- **RMS** is a records management specification, whose definition was led by the National Archives and Records Administration through the Object Management Group.
- The **DOD 5015** is a Department of defense standard focusing on records management.

## Audit and Certification

As highlighted the critical requirement for the public sector, and others like Finance, is compliance, meaning that documentation of what facilities are available, how these relate to legislative requirements and independent verification are essential components of the mix.

The EU is an example of this - In essence Trusted Cloud Europe is exactly this, a program for defining and implementing it across Europe, where they describe a system of best practices, achieved through various methods of certifications of compliance testing.

The EU has begun building their own list of Certifications for this purpose, the CCSL - [Cloud Computing Certification Schemes List](https://www.cloudcomputingcertification.org/), have been mapping interoperability standards, and has also issued [issued RFPs](https://www.cloudcomputingcertification.org/) to begin advancing this area. They anticipate the use of these standards will save them €1 billion a year.

Other relevant documentation from other projects such as the USA have also produced recommendations in this field.

For example the CIO Council’s publication is their own Cloud Best Practices guide: ‘Creating Effective Cloud Computing Contracts for the Federal Government’ - Download here: [cio.gov/cloudbestpractices.pdf](https://cio.gov/cloudbestpractices.pdf)

This describes a best practices framework that addresses much of what the EU calls for and maps to international standards. The guide has a particular emphasis on record keeping due diligence, that could be met through open standards like like [ISO 27037](https://www.iso.org/standard/70540.html) that caters for digital evidence practices.

This further creates operating structure for supplier engagement.

A simple example is [Google announcing their HIPAA compliance](https://www.google.com/hipaa/), and taking from [this Gartner report](https://www.gartner.com/en) on Cloud hosting.
"As of April 2013, Microsoft's “external certification position” can be summarized as:

International Organization for Standardization (ISO) 27001, Standards for Attestation Engagements (SSAE) 16 Standard on Assurance Engagements (ISAE) 3402, EU Model Clauses, and Health Insurance Portability and Accountability Act Business Associate Agreement (HIPAA BAA) certifications apply to Windows Azure core services only: Cloud Services (Web and Worker role instances), storage (Tables, blobs and Queues), virtual networks and virtual machines.

It does not currently cover the rest of Windows Azure features, including Web Sites, SQL Database, Service Bus, Caching, Access Control, CDN, Media Services or Windows Azure Active Directory."

Gartner makes the key point that it does currently cover all of their Cloud services and equally many legal aspects related to Cloud computing are considered somewhat out of date, and so the EU wants to utilize standardization as the driver for this modernization.

A more specific set of guidelines and open standards relevant to EU procurement organizations will create the ‘rising tide that floats all boats', common marketplace arrangements that increase both supplier attractiveness and also ease of aggregated and thus larger procurement by the public sector.

**Procurement Maturity Model**

Further additions planned for this document include a section on the role of SaaS (Software as a Service) and development of a full maturity model to enable use of C-PbD in real-world transformation projects, and will provide tools and templates for Government agencies to integrate the best practices into their RFP procurement practices used to buy Cloud services so they can assess and compare service provider capabilities in these specific areas, via a simple 0,1,2 rating system for none, partial and full compliance.

This sample table provides an initial demonstration of both:

<table>
<thead>
<tr>
<th>Code</th>
<th>Parameter</th>
<th>0 - Not Met</th>
<th>1 - Partially Met</th>
<th>2 - Fully Met</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ability to provide an open API allowing access to its key functionality from external systems.</td>
<td>No API support or lack or support for invoking APIs through internet protocols (ex: Web service, REST, etc.)</td>
<td>APIs supported for providing read-only information. No transactions are supported. Ex: If it’s a HR SaaS system, APIs should be available to deactivate an employee based on</td>
<td>APIs supported for both for read-only and transaction support.</td>
</tr>
</tbody>
</table>
Describe how the hosting is configured (multi-tenancy or single tenancy, multi-instance or single instance).

Supports only single tenant model. No support for scalability.

Supports both single tenant and multi-tenant models. Support for scalability, however it may require manual intervention.

Supports single and multi-tenant at both application and database level (ex: application servers can be common but databases can be different for each tenant). Support for auto scaling based on pre-defined threshold values.

Authors