Trustworthy Computing

A Guide to Data Governance for Privacy, Confidentiality, and Compliance

Part 5: Moving to Cloud Computing

August 2010

Javier Salido, MSc, MBA, CIPP, CIPP/IT
Senior Program Manager, Trustworthy Computing Group, Microsoft Corporation

Doug Cavit, MSc
Principal Security Strategist, Trustworthy Computing Group, Microsoft Corporation

Microsoft Corp. • One Microsoft Way • Redmond, WA 98052-6399 • USA

Contents

Executive Summary

The past decade has seen an unprecedented accumulation of data. Organizations in general and business models in particular increasingly rely on confidential data such as intellectual property, market intelligence, and customers' personal information. Maintaining the privacy and confidentiality of this data, as well as meeting the requirements of a growing list of related compliance obligations, are top concerns for government organizations and enterprises alike. Looking ahead to the coming decade, we can see that with cloud computing, organizations will increasingly have to address the challenges of data protection and compliance. This will require implementing a cross-disciplinary effort within the organization—involving human resources, information technology (IT), legal, and other groups—to devise solutions that address privacy and confidentiality in a holistic way. Data governance is one such approach.

Moving systems and processes into the cloud involves a tradeoff between benefits (such as flexibility, scalability, and a more desirable cost structure) and potential challenges (such as diminished control over data and over the infrastructure that houses and processes that data).

Organizations that want to move confidential data to the cloud should systematically identify incremental risks to data privacy and security in the information lifecycle and risks of noncompliance. For example:

- Organizations should understand the legal implications of moving certain data to a specific cloud services provider (CSP) in a specific geographic location. CSP rights and potential ownership of data should be clarified.
- Organizations should confirm the long-term viability of the CSP, along with the implications and costs of potentially switching to a new CSP in the future.
- Organizations should expect a reasonable level of transparency from the CSP with regard to what security, privacy, and compliance protections are in place. Such transparency should come in the form of documentation, backed up by relevant third-party certifications, about protective measures and processes and how their effectiveness and execution will be verified.
- Organizations should also diagram their data flows to better understand the security, privacy, and compliance risks in the context of the cloud. To do this, they should use the Risk/Gap Analysis Matrix (described later in this paper) to identify gaps in their existing protective and compliance measures and those of the CSP. The matrix helps identify threats and residual risks in specific data flows—valuable information that can be used to improve protections and manage associated risks.

The Whitepaper Series

This whitepaper series aims to answer some key questions that IT managers, security officers, privacy officers, and risk management officers are asking about how to approach the combined challenges of information security and privacy and the associated regulatory compliance obligations.

In its broadest form, data governance is an approach that public and private entities can use to organize one or more aspects of their data management efforts, including business intelligence (BI), data security and privacy, master data management (MDM), and data quality (DQ) management. This series describes the basic elements of a data governance initiative for privacy, confidentiality, and compliance and provides practical guidance to help organizations get started down this path.

The series is meant for organizations of all sizes, including those with regional as well as global focus and those with on-premises systems as well as cloud-based systems. Some might already have an effective IT governance process and information security management system (ISMS) in place, as well as successful privacy and risk management efforts. Some might just be getting started.

At Microsoft, we believe that in order to deal effectively and efficiently with data confidentiality and privacy challenges, organizations must adopt a proactive stance, one in which they hold themselves accountable for:

- Appropriately protecting the security of customers' and employees' personal information, as well as the organization's intellectual property and trade secrets
- Respecting, preserving, and enforcing customer choice and consent throughout the information lifecycle, particularly when it comes to deciding how personal information is used and distributed within and outside the organization

In approaching data privacy and security, organizations should also consider the following:

- Taking a holistic approach to data privacy and security needs as well as related regulatory and internal compliance requirements. This approach to the planning and implementation of data privacy and security brings together a range of participants. They could include groups and individuals that:
    - Own business processes that generate, collect, and use data
    - Have specific charters with respect to confidential data, such as the chief privacy officer, the legal department, and the IT department
- Augmenting approaches that focus on mere compliance with "the letter of the law" by implementing and enforcing data privacy and security measures based on generally accepted principles,[1] state-of-the-art industry best practices, and self-regulation measures that go beyond mere compliance with regulations and standards.

[1] Organisation for Economic Co-operation and Development, "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data," www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html. American Institute of Certified Public Accountants (AICPA) and Canadian Institute of Chartered Accountants (CICA), "Generally Accepted Privacy Principles," www.aicpa.org/InterestAreas/InformationTechnology/Resources/DataIntegration/DownloadableDocuments/GAPP_%20Practitioner_%20092006.pdf.

- Augmenting prevailing IT privacy and security paradigms—which address threats by restricting access to data and keeping it from "escaping" well-defined boundaries[2]—by evaluating threats to confidential data at different stages of the information lifecycle. This approach helps organizations identify technical and nontechnical measures that can reduce security and privacy risks to acceptable levels.

The first paper in the series analyzes the challenges that organizations face today when trying to protect data privacy and security in on-premises systems, including an increasingly complex regulatory environment. It also looks at the concept of data governance and how it can complement ongoing efforts within the organization.

The second paper looks at two of the core capability areas that an organization must develop as part of a data governance for privacy, confidentiality, and compliance (DGPC) initiative: People and Process.

The third paper analyzes the last of the three core capability areas, Technology. It discusses a risk analysis process and an associated threat modeling technique that can help organizations identify threats against data security and privacy, as well as threats arising from noncompliance in specific data flows, and manage the associated risks.

The fourth paper offers a capability maturity model that can help organizations identify their level of DGPC maturity in the People, Process, and Technology core capability areas, determine appropriate target maturity levels, and develop action plans to reach those targets.

This paper, the fifth in the series, discusses how DGPC challenges change when organizations opt to use cloud-based services, and how the techniques discussed in the third paper can be applied in the new context.

[2] Weitzner, Abelson, Berners-Lee, Feigenbaum, Hendler, and Sussman, "Information Accountability," Massachusetts Institute of Technology CSAIL Technical Report, June 2007, http://dspace.mit.edu/bitstream/handle/1721.1/37600/MIT-CSAIL-TR-2007-034.pdf?sequence=2.

An Introduction to Cloud Computing

The term *cloud computing* refers to several different computing paradigms, not all of which are completely new. In fact, opinions differ on what qualifies as a "cloud service" and how these services should be classified. The National Institute of Standards and Technology (NIST) in the United States has proposed three cloud computing *service models*:

- Software as a Service (SaaS), in which software applications are provided and managed in the cloud by a Cloud Service Provider (CSP). An example of this is Microsoft® Online Services, which offer hosted versions of Microsoft Exchange and Microsoft SharePoint® for use by both public and private organizations.[3]
- Platform as a Service (PaaS), in which a CSP delivers the underlying infrastructure, including operating systems and storage, and allows organizations to build and run applications using languages and tools provided and supported by the CSP. An example is Microsoft's Windows Azure™ platform.[4]
- Infrastructure as a Service (IaaS), in which a CSP gives an organization access to basic IT infrastructure (network, hardware, core operating system, and virtualization software) on which the organization can deploy its own applications and data in a virtualized environment, applications that were developed using languages and tools that are not provided or supported by the CSP. Examples include Amazon's EC2 and Rackspace's Cloud Servers.

Key characteristics of all three models include convenience, cost effectiveness, flexibility, and elasticity—all made possible through the sharing of resources by multiple tenants and by rapid provisioning through self-service—including adding additional processing power and storage as needed.

Cloud *deployment models* can take the form of private clouds operated solely by or for a single organization; community clouds for groups of organizations with similar service requirements; and public clouds with one general service-level agreement (SLA) that applies to all customers, in which data resides on shared resources. In some limited scenarios, hybrid clouds can be built by connecting public and private clouds and sharing services and data among them.

Another key distinction of cloud computing is that information storage and processing need not be limited by space or geography. Indeed, cloud users typically don't need to know how many "virtual filing boxes" they will use because the available storage and processing power can scale to meet their needs.

[3] www.microsoft.com/online/products.mspx

[4] www.microsoft.com/windowsazure

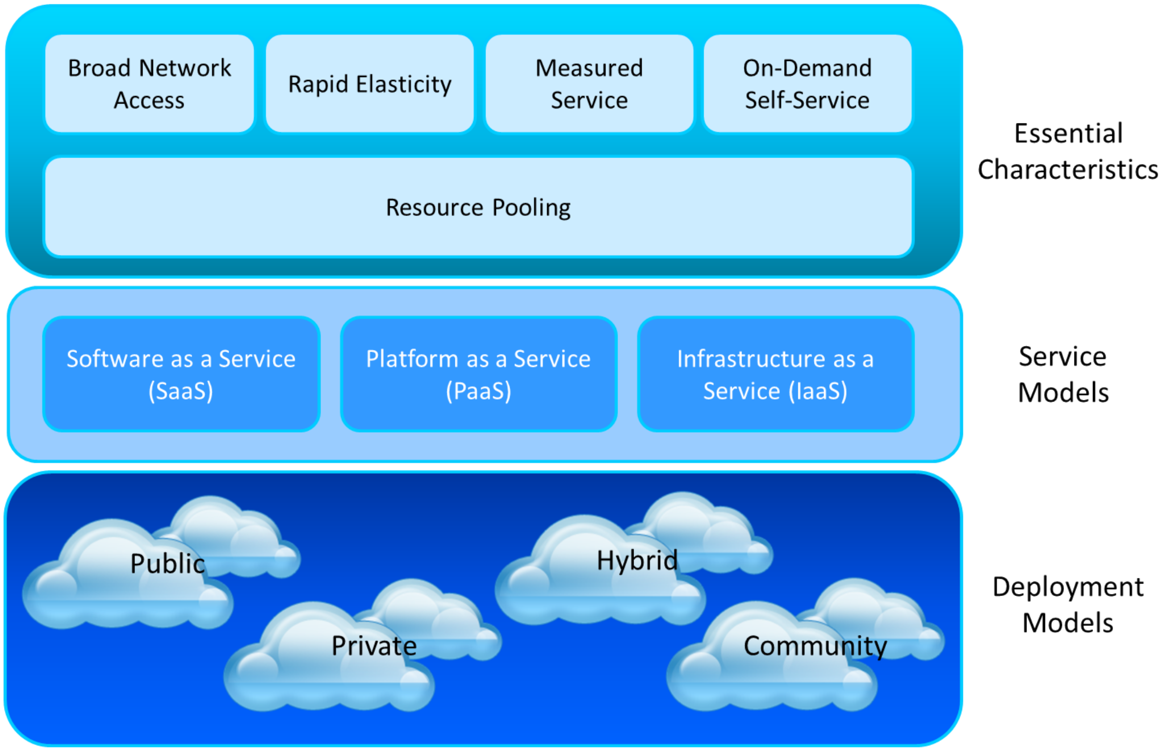


Figure 1. Depiction of the NIST working definition of cloud computing.[5]

Off-premises cloud services offer many potential advantages, including security improvements, flexible scaling, and reduced or no capital spending on IT. However, tradeoffs are inevitable, and organizations should carefully consider these in their business planning and risk management planning.

In an on-premises model, the organization is responsible for all aspects of IT—people, processes, and technology. The organization buys the hardware, licenses the software, secures the datacenters, defines processes and procedures, and hires the people who run everything.

As a result, the organization is responsible for, and has control over, the physical location of the datacenter (which determines key factors such as which country's laws apply); the security of the datacenter; how employee, partner, and customer personal information is used and distributed; the trustworthiness of system administrators; and the documented information security program that protects the confidentiality, integrity, and availability of data and systems (including, but not limited to, configuration, patching, incident response, and business continuity management).

[5] http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc

By contrast, particularly in non-private clouds, many of these functions may be handled by the CSP, whose system administrators might be its own employees or even third parties it has hired. These types of arrangements have some parallels with common scenarios in which organizations outsource critical IT functions to third parties. But some of the legal, operational, and security-related complexities are unique to cloud services and require special attention.

Many elements of cloud services represent wholesale change. For example, to make cloud services capable of expanding flexibly, hardware is often shared among customers, and the "security boundary" between them may be virtual (through the use of virtualized compartments) rather than physical (through the use of separate hardware). In addition, on-the-fly allocation of extra resources might mean that the geographic location of data depends on scalability, availability, or other factors rather than on security and jurisdictional considerations, especially when a CSP has datacenters in multiple jurisdictions. This can create uncertainty about which laws apply to the handling of the data.

The following sections will address these and other considerations, as well as provide guidance on how to approach them.

What Changes in the Cloud

As with other technological shifts, moving to cloud computing requires organizations to address privacy and security considerations—including compliance and risk management, identity and access management, service integrity, endpoint integrity, and information protection.

When an organization shifts to using cloud services, its IT department might have to adapt to a data management regime and practices that are not under their direct control. This is especially true when an organization opts to use a combination of on-premises and cloud systems, which might mean that new and extended security processes must encompass multiple providers to achieve comprehensive protection of information. Risk management and privacy and security management remain the responsibility of the organization and must be extended to include the CSP.[6]

[6] These are just some of the important considerations. For further details, see the guidance on cloud computing from the Cloud Security Alliance (CSA) at www.cloudsecurityalliance.org and the European Network and Security Agency (ENISA) at www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment.

Information Protection

The sensitivity of the data involved in the use of a service is critical to determining whether the service can be managed by a CSP and, if so, which access controls can and should be used to ensure that compliance obligations are met throughout the process. Factors that must be considered include the following:

Data classification. Defining and systematically adhering to a sound data classification policy—for instance, specifying which types of data are considered confidential and which are not—is critical to determining the control mechanisms that will protect each data type. While this principle also applies to on-premises systems, risks derived from having no data classification policy or one that is incorrect are greater in the cloud because data might not be afforded the appropriate protective measures. There is little the CSP can do in this area; the responsibility mostly lies with the organization that is moving data to the cloud.

Organizations should implement a **data classification policy and procedures** for deciding which data is ready for the cloud, under which circumstances, and using which controls.

Data quality. When data is collected or processed in the cloud, consistent data definitions are needed to maintain the quality of service to customers, limit costs related to data cleansing and cloud resource consumption, and facilitate data classification. Maintaining the quality of employee and customer personal data is a widely accepted privacy principle because low-quality data can lead to situations in which individuals receive communications that they opted not to receive or they fail to receive communications that they should have received. Low-quality data can also result in customers being denied services that they should have received.[7] Another challenge is that of maintaining cloud stored data in synchronization with on-premises copies of the same data.

Protective measures. When the management and control of information moves to the CSP, the customer organization's ability to protect, retrieve, or move information is reduced. It is therefore important for the organization to understand who controls the identity and authorization system for access to information, where backup data is stored, whether data encryption is supported, what cost is associated with the encryption solution (e.g., feature loss), and so on.

Data partitioning and processing. If data is stored in a "public cloud," it might reside on infrastructure shared with other organizations' data. Strong data protection practices should be followed to ensure that data is partitioned and processed appropriately. Organizations should understand who has access to their data and consider whether that risk is acceptable before they allow their data to be processed in the cloud. They must also understand the architecture of the CSP and gain assurances that shared virtual machines are secured against attacks from other virtual machines on the same physical hardware.

[7] Organisation for Economic Co-operation and Development, "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data," www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html.

Finally, certain legal aspects of the agreement between the organization and the CSP might affect information protection—for example, how access to data will be granted and managed if there is a dispute with the service provider. Can the CSP guarantee that it will not retain the organization's data if they cancel the service?

Compliance and Risk Management

With an on-premises computing system, organizations have primary control over how the environment is built and run. In a cloud scenario, some of the related tasks and decisions are delegated to the CSP. This can present new challenges, such as the need to entrust parts of the organization's fundamental compliance and risk management processes to the CSP.

Delegation does not discharge the organization from managing risk and compliance, or from having to prove compliance to the appropriate authorities. In fact, CSPs generally exclude themselves from compliance responsibility in their service agreements. Nevertheless, the CSP's controls might help the organization in its compliance efforts.

Compliance requirements can be fulfilled by a **skilled internal team** and a certain level of **process transparency** on the part of the cloud service provider.

Appropriately managing risk and compliance requires some transparency on the part of the CSP about its operations. At the same time, CSPs have a responsibility to prevent their customers from compromising the environment shared by multiple parties and therefore must limit disclosure of certain security-related details.

Thus, the resulting agreement between the customer organization and the CSP might include a combination of contractual commitments, an agreed-upon level of visibility into the CSP's control framework, third-party verified certifications and attestations, and other elements that the customer organization can integrate into its risk and compliance management environment through its own data governance program.

Identity and Access Management

Secure identity management is essential in any environment, but it can become more challenging in a cloud computing context—particularly when the requirements include provisioning, de-provisioning, and managing user accounts as well as managing access to confidential data resources. Cloud systems might need to cross the boundaries of multiple departments, organizations, and CSPs that use different products, technologies, and processes.

Any digital identity system for the cloud must be **interoperable** across organizations and cloud providers and must be managed through strong processes.

As a result, cloud-based services need highly secure cross-domain collaboration, with protections against the misuse of the identities of people and devices. Identity and access control, especially for higher-value assets, should be based on robust cryptographic credentials that enable a "claims-based" system that authenticates the claims made by any trusted entity.[8] The strength of that authentication system should be reasonably balanced with the need to protect the privacy of users. To help achieve this balance, the identity and access management system should allow strong claims to be transmitted and verified, without revealing more information than is necessary for any given transaction or connection within the service. As a general rule, the less information the system reveals, the lower its risk of improperly exposing personal information.

On the process side, organizations should understand how the processes for provisioning, de-provisioning, and managing accounts and access rights will work with a specific CSP. Particularly in complex environments with multiple claims providers, it is critical to establish which organization will "own" the identities of users, what controls will surround identity and access management, how identity federation will work with different providers, and how ad hoc collaboration with people outside the organization will be handled. The cost and implications of a potential future switch to different "claims providers" should also be carefully considered.

Any identity environment that extends to the cloud must be interoperable across applications that consume identity claims, must enable the secure migration of data access controls to the cloud and back, and must be manageable for the organization that is paying for the service.

Service Integrity

Service integrity has two components: 1) service engineering and development and 2) service delivery. Service engineering and development refers to how the CSP incorporates security and privacy by design into all phases of development and deployment. Service delivery refers to the way the service is operated to meet contractual levels of reliability and support.

Service Engineering and Development

Any organization that develops software should follow an internally transparent engineering and development process that incorporates security and privacy into its products by design. Engineering and development for cloud computing environments—whether by a customer organization's development group or by the CSP and/or third parties—is no different in this regard. Microsoft uses its Security Development Lifecycle (SDL) and Privacy Guidelines for Developing Software Products and Services in the development of all Microsoft software products and cloud services, and it makes these guidelines available to the general public.[9]

The CSP should follow a **clear, defined, and provable process** to integrate security and privacy, by design, into the service from development through operations.

When evaluating a CSP, organizations should ask how the provider builds security and privacy into the development process, how often threat models are created and updated, how the effectiveness of the security

[8] www.identityblog.com/wp-content/images/2009/06/UserCentricIdentityMetasystem.pdf

[9] Security Development Lifecycle: http://msdn.microsoft.com/en-us/library/ms995349.aspx. Privacy Guidelines for Developing Software Products and Services: www.microsoft.com/privacy/guidance.aspx.

and privacy response group is measured, and how customers are informed about security updates and automatic update deployment plans.

Service Delivery

Before migrating business-critical processes to the cloud, an organization should have clearly defined privacy and security processes, and it might need to make changes before engaging with one or more CSPs. It might, for example, need to review its security monitoring, auditing, forensics, incident response, data breach notification, and business continuity processes as well as its enforcement of choice and consent associated with personal information. It might also need to review and update key policies such as its privacy policy and data classification policy. These changes should be defined before the initial cloud delivery setup, taking both parties' needs and capabilities into account.

For certain applications or services, privacy and security requirements might be fairly simple and straightforward. For other services, such as those involving highly confidential assets, more stringent requirements might need to be established for physical security, logging, background checks on administrators, acceptable geographic locations, access control, usage limitations, and so forth.

The **service delivery capabilities of the CSP** and the privacy and security management and auditing **needs of the customer** should be aligned.

Cloud systems that store and process confidential information and other high-value assets might require detailed plans for managing performance problems, including conducting network and image forensics. Services should include response contacts and restoration processes in the event of interruption in service delivery. Also, agreements should define which security monitoring and auditing capabilities will be provided by the CSP and at what price.

Endpoint Integrity

Discussions about cloud security and privacy often focus on the service itself, as well as on the CSP's privacy and security quality and practices. But it is essential to evaluate the entire service chain and the data flows throughout the data lifecycle in order to avoid flaws in service design and delivery. Cloud services begin and end either within an organization or at the PC or device of the individual using the service. When data privacy or security is compromised or a data breach occurs, the issues often occur on individual PCs or even storage media, not on the backend servers. To increase the end-to-end trustworthiness of cloud computing, organizations must consider the full spectrum of activity to help protect customers, partners, and employees from threats (such as online identity theft, Web site cross-scripting attacks, phishing attacks, malicious software downloads, and theft of endpoint devices or storage media).

It is important to **include the endpoint** when considering privacy and security in cloud-based services.

Many organizations have internal risk management programs to help protect the endpoints and manage information privacy and security. Before an organization moves to a cloud environment however, it should review each cloud service's privacy and security measures. It is particularly important to note that some solutions integrate services from multiple CSPs with varying levels of visibility. The organization should identify and address any existing dependencies before signing on with a service.

If the CSP is responsible for endpoint security—a practice known as Security as a Service—the important questions to ask should include: How are security, privacy, and compliance requirements enforced? How is data protected against theft, loss, unauthorized distribution, or misuse? Can the customer still use encryption or a rights management mechanism to protect data from loss or theft? Can the service be restricted to specific authorized endpoints or machines? The customer organization should also implement a training program for its internal developers, IT personnel, and end users to explain how to protect the organization's data and systems in the cloud environment.

Elements to Consider When Moving to the Cloud

Our discussion has focused thus far on elements that can change when an organization moves on-premises systems and data to the cloud. This section discusses key questions that organizations should ask before choosing a service model, a type of cloud service, or a CSP. Keep in mind that data quality and availability are considered important components of data privacy.

Viability of the CSP and Potential Switching Costs

The issue of a CSP's viability has many parallels with what organizations have dealt with for years in vendor outsourcing agreements. The core issues of financial stability, capital investment, basic service guarantees, and avenues of recourse if the service agreement is breached are all relevant to cloud services. However, the cloud scenario introduces other issues relating to multi-tenancy and switching providers:

- The customer organization should obtain guarantees about bandwidth and availability in multi-tenant environments. The CSP should set clear expectations about scalability and protections against service disruptions due to scaling of activities by other customers.
- The customer organization should find out whether it can switch providers if the CSP fails to meet expectations, and what the switching costs would be.
- Other issues related to switching providers include:
  - Retaining ownership of domain names.
  - Data portability.
  - Application portability, particularly in a PaaS scenario, and associated costs.
  - The cost of data migration to a different service, especially one with very different facilities for hosting important databases.
  - Portability of identity and access controls and associated costs. Many CSPs expect the customer to use the CSP's identity and access control system. If the organization wants to move to a different provider, it might be forced to re-provision all those user accounts. This argues for implementing an identity claims metasystem in which the customer organization manages identity and access control rather than outsourcing that function to the CSP.

Transparency

Organizations must be confident that their CSP takes the threat of malicious attacks seriously, that it will make reasonable efforts to protect data entrusted to them from thieves and hackers, and that it will minimize potential attack surfaces. Organizations must also be confident that the CSP will abide by known, unambiguous rules governing how customers' and employees' information is processed, used, stored, and possibly distributed to and shared with third parties.

Processes and Procedures

A trusting relationship between the customer organization and the CSP requires that the CSP adhere to appropriate privacy and security practices from software development to service delivery, operation, and support. CSPs should back up their claims of having effective protective mechanisms in place. They must provide reasonable disclosure of those mechanisms; the standards, principles, and industry best practices they are based on; and how their effectiveness is verified.

To gain these assurances, organizations should ask the CSP for evidence that it maintains a comprehensive and properly documented information privacy and security program—one that is kept updated and provides safeguards appropriate to the organization's needs.[10] The CSP might respond by providing a high-level summary that includes items such as the following:

- Whether its information privacy and security program complies with leading industry standards, such as the Payment Card Industry Data Security Standard (PCI DSS), ISO/IEC 27001:2005 (Information Security Management Systems – Requirements), and NIST SP800-53 under the U.S. Federal Information Security Management Act. The CSP should provide independently verified certification and attestation documents.
- Whether it uses appropriately robust authentication and authorization mechanisms, what those are, and how processes such as account provisioning and management, along with access control, are managed and by whom.
- Whether third parties and contractors participate in activities related to datacenter and data management or any other activities that could grant them access to the organization's data, what roles they play, and whether and how they are vetted.
- Whether applications and other components of the service (both hardware and software) receive thorough security testing before deployment.

[10] For real-life examples of such documentation, see "Microsoft's Compliance Framework for Online Services" and "Securing Microsoft's Cloud Infrastructure" at www.globalfoundationservices.com/security/index.html. See also www.microsoft.com/downloads/details.aspx?FamilyID=5736aaac-994c-4410-b7ce-bdea505a3413&displaylang=en.

- Whether and how software design, development, and deployment practices and processes, particularly in the IaaS and SaaS models, consider security and privacy needs and whether the technologies and features meet security and privacy requirements. Transparent and widely accepted practices such as Microsoft's Security Development Lifecycle and the associated Privacy Guidelines for Developing Software Products and Services can serve as important proof points.[11]

Policies

Organizations should ask the CSP about its policies and practices that affect the ownership, use, and retention of data (or related aggregated data and metadata) that is stored with the CSP. Questions should include:

- What policies and practices, including those related to data availability, will determine the geographic location of the customer organization's data?
- What are the CSP's policies and practices relating to data ownership and retention?
- Can the CSP guarantee partial or total destruction of the data in its safekeeping, including mirrored and backup data, at the request of the customer organization?
- What are the CSP's policies and practices regarding software updates by the CSP that might affect the customer organization's operations?
- For multi-tenant services, what are the CSP's policies governing the assignment of tenants and the sharing of logical and physical resources?

Compliance and Related Issues

When thinking about compliance obligations related to data stored and processed in the cloud, organizations should consider two issues. First, should the data in question actually be placed in the cloud, and what conditions would have to be met to do this? Second, what assistance can the CSP provide to help the organization meet the applicable compliance obligations? To answer these questions, the organization must first understand what its own regulatory and internal policy requirements are. Microsoft recommends that organizations develop a data classification policy and a set of "harmonized compliance requirements"—a list that summarizes the organization's regulatory and internal policy obligations and that can, in turn, be used to define the requirements the organization has for the CSP.[12]

[11] Security Development Lifecycle, http://msdn.microsoft.com/en-us/library/ms995349.aspx. Privacy Guidelines for Developing Software Products and Services, www.microsoft.com/privacy/guidance.aspx.

[12] See the discussion of the Process core capability area of the DGPC framework in "A Guide to Data Governance for Privacy, Confidentiality, and Compliance: Part 2: People and Process" at www.microsoft.com/datagovernance.

Processing and Hosting Data in the Cloud or with a Specific CSP

The applicable legal framework might vary depending on factors such as the organization's vertical industry and the geographic location of the CSP, the datacenter, or the organization itself. Therefore, laws and regulations will likely play an important role in determining whether an organization should store specific types of data with a specific CSP. As an example, but by no means the only case, if the organization is based in the European Union, transborder flows of personal data, particularly outside the European Economic Area (EEA) are subject to significant restrictions.

Another example would be that based on its analysis of the applicable legal framework and its own industry's circumstances, the organization might want to consider working with the CSP to implement private cloud services.

Equally important are the terms of the agreement with the CSP:

- For what purposes may the data be used? This might seem obvious, but the terms of the agreement might, for instance, allow the CSP to collect, process, and use the organization's data, transactional data, aggregated data, or metadata for purposes (such as market research, customer profiling, or a case study to publicize the CSP's offerings) that are not directly related to the specific services it provides to the organization.[13]
- To what extent is the CSP responsible for mistakes and lapses on the part of third parties and contractors that it employs?
- How will the CSP address key issues such as data retention and data destruction?

Meeting Compliance Obligations

Once an organization establishes the viability of storing and processing data in the cloud, it can address how to meet its regulatory compliance obligations as well as data privacy and security-related commitments it has made to customers, shareholders, employees, and other stakeholders. Just as important is the question of how it will show proof of compliance. Given the decreased control it has over data and over protective measures and reporting capabilities in the cloud, the organization must coordinate with the CSP. The CSP should clarify the processes and escalation paths it will follow in exceptional circumstances, such as notifying the organization in case of a data breach.

The terms of agreement should include a list of compliance-related documents that will be provided by the CSP—including certifications, plans, and escalation paths—preferably along with relevant templates or examples.

To understand and manage the risks associated with moving confidential data to the cloud, organizations should implement a risk-based approach. The following section will discuss a technique that organizations of all types and sizes can use to implement such an approach in the context of their existing data governance program.[14]

[13] "Privacy in the Clouds: Risks to Privacy and Confidentiality from Cloud Computing," Robert Gellman, World Privacy Forum, www.worldprivacyforum.org/cloudprivacy.html.

[14] For more information on data governance, see the other papers in this series at www.microsoft.com/datagovernance.

Integrating the Cloud into the DGPC Framework

The DGPC framework's Technology core capability area focuses on selecting technical and manual controls to keep security, privacy, and compliance risks to an acceptable level. This approach involves filling out a Risk/Gap Analysis Matrix built around three elements: the information lifecycle, the four technology domains, and the organization's data privacy and confidentiality principles. The matrix (described in more detail later in the paper) is useful for analyzing data flows in any environment, including ones that combine on-premises and cloud-based systems.

Information Lifecycle

To select appropriate technical controls and activities to protect confidential data, an organization must first understand how information flows over time and how it is accessed and processed at different stages—by multiple applications and people, and for various purposes. Figure 2 illustrates the concept of the information lifecycle. Understanding the risks within each stage helps clarify what safeguards are needed to mitigate those risks.

Most IT professionals are well acquainted with these lifecycle stages,[15] so we will not discuss them in detail here except for one important aspect: the need to recognize a Transfer stage. As data is copied or removed from storage as part of a transfer to a new system or data flow, a new information lifecycle begins. Organizations should place as much emphasis on the security and privacy of data that is being transferred to a different location (typically a new system) as they do for the original dataset. In the cloud, this requires understanding key aspects of the transfer vehicle—the Internet—and its risks. It also requires understanding how the CSP's policies, systems, and practices might differ from those of the organization that collects the data.

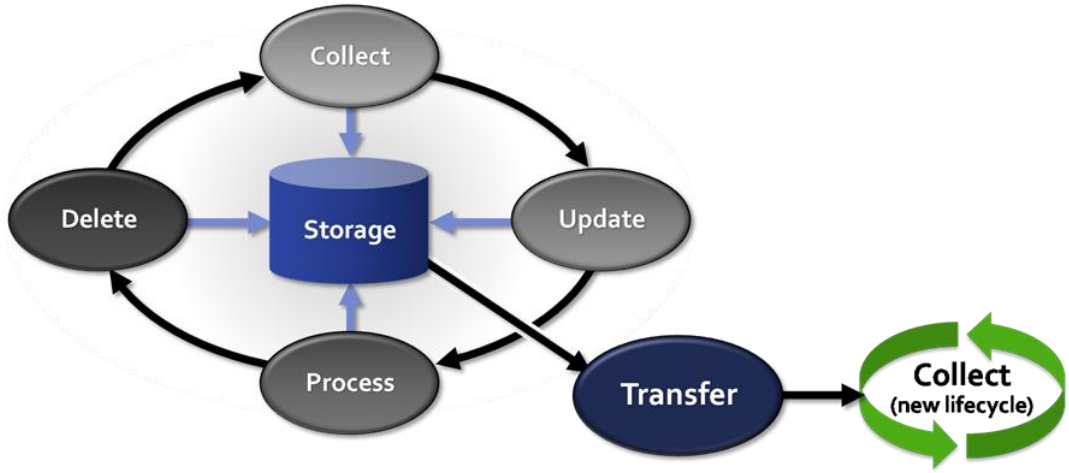


Figure 2. The information lifecycle.

[15] For a more detailed discussion of the information lifecycle, see "A Guide to Data Governance for Privacy, Confidentiality, and Compliance: Part 3: Managing Technological Risk" at www.microsoft.com/datagovernance.

Technology Domains

The Risk/Gap Analysis Matrix will also address the four core technology domains that must be considered when evaluating the effectiveness of data protection in a cloud services model:

- **Secure Infrastructure.** Infrastructure security requires looking at the entire technology stack in a holistic way and at each level to understand the CSP's policies for building and maintaining the infrastructure in a secure manner. Organizations should ask the CSP to provide details about the entire technology stack, including but not limited to:
    - The physical security and mechanical robustness of the datacenters
    - Controls used to commission and decommission equipment within the datacenter, including hardware security controls such as TPM chips or hardware encryption devices
    - Network operations and security features, including firewalls, protection against distributed denial of service (DDoS) attacks, integrity, file/log management, and anti-virus protection
    - Basic IT controls and policies governing personnel, access, notification of administrator intervention, levels of access, and logging of access events
- **Identity and Access Control.** Identity and access control is among the most overlooked and difficult IT tasks, even though it has the most direct consequences in terms of information protection. It involves the following components:
    - **Identity provisioning.** The CSP should integrate its IT practices with those of the organization so no security gaps exist when it comes to provisioning new users, creating trust relationships for access control, and de-provisioning users whose status has changed.
    - **Authentication.** The CSP should support different levels of authentication depending on the customer perception of the nature of the service and the sensitivity of the data entrusted to the service.
    - **Identity federation.** Identity federation can take multiple forms. In some cases, the customer organization is required to use the CSP's identity infrastructure or to employ very simple mechanisms for authenticating users from different identity domains. However, the method with the greatest potential to enhance privacy and that offers the most flexibility is an identity claims model (as implemented in Microsoft Azure, Identity Federation Server 2.0, and other advanced identity mechanisms built on SAML 2.0). With this type of mechanism, the customer organization maintains complete ownership and control of business-critical portions of the access control stack. For example, it can maintain control of identity (provisioning and de-provisioning of accounts), authentication, and authorization while outsourcing access control to the CSP. Flexibility lies in the customer organization being able to establish trust relationships with other organizations, as opposed to having to include the CSP in the relationship. This allows the identity claims model to provide differing levels of access control that are independent of the CSP or any other entity in the trust ecosystem. This also makes sensitive data much more easily portable because key parts of the access control decision are no longer tied to the particular CSP.

  - **Standards.** To achieve the requisite level of federation and application portability, organizations should evaluate the CSP's adherence to industry standards governing identity, authentication, authorization, and access. An example of these standards is SAML 2.0 tokens.
  - **Auditability.** All access-control decision points should be auditable so it is easy to identify unauthorized access, and hold unauthorized users accountable. This includes unauthorized access by means of administrative credentials maintained by the CSP.
- **Information Protection.** The requirements in this area will depend on how critical the data is and the type of service used.
  - **Data confidentiality.** Whenever possible, confidential data should be encrypted (and decrypted) during on-premises or end-point processing before it is transferred to the cloud. The key concern is protecting data confidentiality in an end-to-end fashion.
  - **Basic data integrity.** Key concerns in this area include infrastructure reliability, access controls, and commingling of data.
  - **Data availability.** Service availability requirements should be defined. In addition, if data becomes corrupted, alternative storage, backup, or other mechanisms should be available to protect the information.
  - **Data persistence.** Issues of data persistence include making backups, maintaining multiple copies, and using virtual machines. Issues of forensic availability for civil or criminal law enforcement should also be addressed. It is prudent to include a data persistence review in reviews or audits of data retention policies and procedures.
- **Auditing and Reporting.** Auditing and reporting are the keys to understanding what happens to data that is not under the organization's direct control. Without them, it is difficult to roll back unwanted or fraudulent transactions. Auditing also forms the basis for compliance regimes. Here are the main concerns in this area:
  - **Audit scope.** What exactly is audited in the service? How comprehensive are the audits, and how long does audit information persist? Is user information persisted for forensic analysis? Can audit information be used to roll back improper transactions? Do audits conform to relevant laws, regulations, standards, and industry best practices?
  - **Audit integrity.** How is audit information protected? Who has administrative access to it? Is the audit information stored in a protected and reliable manner?
  - **Reporting.** Is the audit information easily accessible? Does it have sufficient scope for compliance and governance controls? Is the information usable as a forensic artifact for legal purposes?

Data Privacy and Confidentiality Principles

The final element of the Risk/Gap Analysis Matrix is comprised of the four guiding principles described below, which can help organizations identify gaps in existing protective measures and select (or request from the CSP) technologies and activities that will protect their confidential data. Each high-level statement should be followed by more detailed guidance to help the data governance team identify specific threats. (The detailed guidance should be customized to the specific needs of the organization.[16])

- **Principle 1: Honor policies throughout the confidential data lifespan.**[17] This includes a commitment to process all data in accordance with applicable statutes and regulations, preserve privacy, respect user choice and consent, and allow individuals to review and correct their information if necessary. This principle encompasses the following concerns:
  - Choice and consent regarding the collection, use, and disclosure of customers' and employees' personal information
    - Provide adequate notice of data collection, use, disclosure, and redress policies.
    - Use clear and unambiguous language when informing users about their choices and request consent for the collection and use of personal information.
  - Individual access to personal information
    - Provide clear and simple ways for users to review and correct their personal information.
  - Accountability
    - Implement adequate controls to classify information, enforce user choice and consent, retain and destroy data, and meet compliance obligations.
- **Principle 2: Minimize risk of unauthorized access or misuse of confidential data.** The information security management system (ISMS) should provide reasonable administrative, technical, and physical safeguards to ensure the confidentiality, integrity, and availability of data. This principle encompasses the following concerns:
  - Information protection
    - Protect structured and unstructured data while it is in transit or at rest.
    - Limit the amount of information that can be obtained from the cloud by a single user in a single operation and/or in a single day.
    - Define and enforce clear account and access provisioning controls.
  - Data quality
    - Implement methods to verify the accuracy and timeliness of data.
    - Allow users to make corrections as appropriate.
- **Principle 3: Minimize the impact of confidential data loss.** Information protection systems should provide reasonable safeguards, such as encryption, to ensure the confidentiality of data that is

[16] For a questionnaire that can be useful for customization, see "Application Privacy Assessment" at www.microsoft.com/datagovernance.

[17] These policies might consist of requirements derived from laws, standards, promises, individual customer or employee choices, commercial obligations, and other sources.

lost or stolen. Appropriate response plans and escalation paths should be in place for critical incidents, and all employees likely to be involved should receive adequate training. Key concerns include:

- Information protection
  - Provide mechanisms to ensure that information is useless to other parties if it is stolen or lost.
- Accountability
  - Implement a data breach response plan and define an escalation path.
  - Verify adherence to data protection principles through appropriate monitoring, auditing, and use of controls.
  - Leverage monitoring and auditing capabilities to hold transgressors accountable for policy violations.

- **Principle 4: Document applicable controls and demonstrate their effectiveness.** To help ensure accountability, the organization should verify adherence to data privacy and confidentiality principles through appropriate monitoring, auditing, and use of controls. This principle encompasses the following concerns:
  - Accountability
    - Properly document plans, controls, processes, and system configurations.
  - Compliance
    - Ensure that compliance can be verified and demonstrated through existing logs, reports, controls, and process documentation.
    - Define a clear noncompliance escalation path and process.

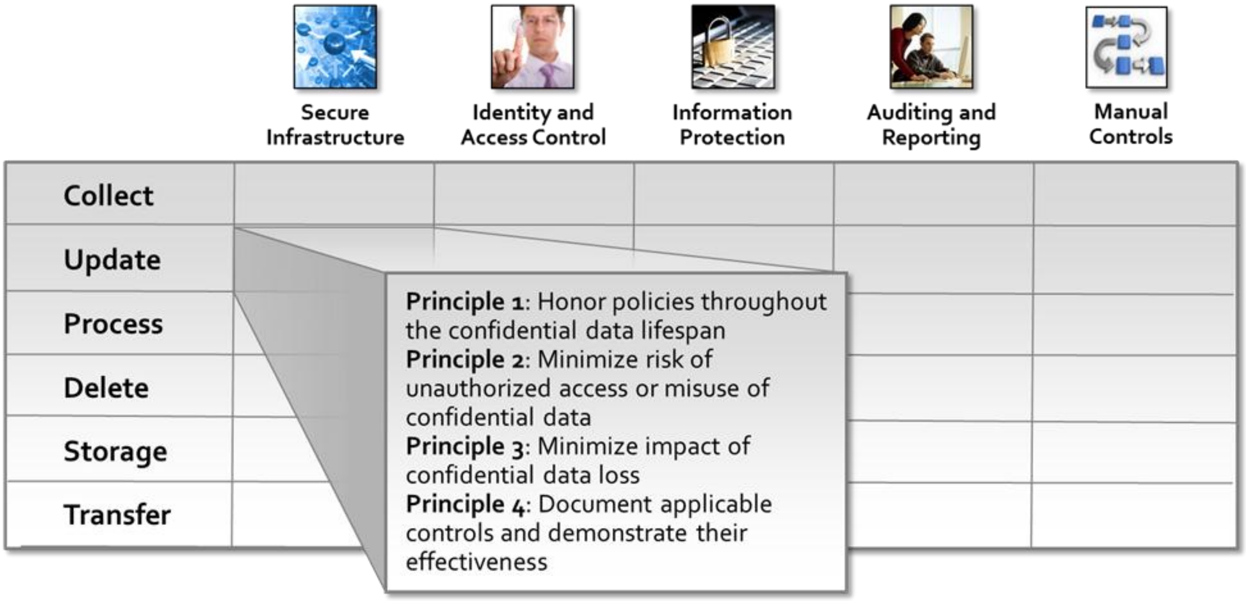
The Risk/Gap Analysis Matrix

The Risk/Gap Analysis Matrix, shown in Figure 3, brings together the information lifecycle, technology domains, and data privacy and confidentiality principles in a tool that helps organizations identify and address gaps in their efforts to protect data against privacy, confidentiality, and compliance threats within specific data flows. Use of the matrix provides a unified view of existing and proposed protection technologies, measures, and activities.

Each row depicts a stage of the information lifecycle. The first four columns in the matrix each represent a technology domain, while the far-right column represents manual control activities that must take place to meet the requirements of the four data privacy and confidentiality principles at each stage of the information lifecycle. The cells of the matrix essentially translate the four principles into questions:

- Is the system honoring the policies of the organization (and meeting compliance obligations) at each stage of the information lifecycle and for each column in the matrix?
- Is the system minimizing the risk of unauthorized access or misuse of confidential data?
- Is the system minimizing the impact of confidential data loss?
- Is the system documenting all applicable controls? Can their application and effectiveness be demonstrated?

In the case of data stored and processed in the cloud, the level of detailed knowledge that an organization will have about the CSP's protective measures and activities will likely be much less than for on-premises systems. Nevertheless, this matrix can help an organization evaluate and understand potential risks, make informed decisions about CSPs and their levels of transparency and data protection, and make inquiries about their protective measures.



Principle 1: Honor policies throughout the confidential data lifespan
Principle 2: Minimize risk of unauthorized access or misuse of confidential data
Principle 3: Minimize impact of confidential data loss
Principle 4: Document applicable controls and demonstrate their effectiveness

Figure 3. The Risk/Gap Analysis Matrix.

Assessing Risks Using the Risk/Gap Analysis Matrix

The matrix is a powerful tool for the risk assessment and mitigation process, which has five stages, as shown in Figure 4 and described below.

Step 1: Establishing the risk analysis context. This step involves defining the business purpose of the data flow, defining the use cases (how the data will be used and what systems will be involved), and identifying the privacy, security, and compliance objectives for the flow.

Step 2: Threat modeling. This step involves diagramming—creating a graphical representation of—the data flow and enumerating potential threats against privacy, security, and compliance that could affect the data flow.

Microsoft's product teams and consulting services organization typically employ data flow diagrams (DFDs) with the addition of "trust boundaries." A trust boundary is a border that separates business entities and/or IT infrastructure realms, such as networks or administrative domains. Every time confidential data crosses a trust boundary, assumptions about security, policies, processes, and practices—or all of these combined—that were valid for one entity may change, and with them the threats that might affect the flow. A detailed description of DFDs and trust boundaries can be found in the Microsoft IT Infrastructure Threat Modeling Guide.[18]

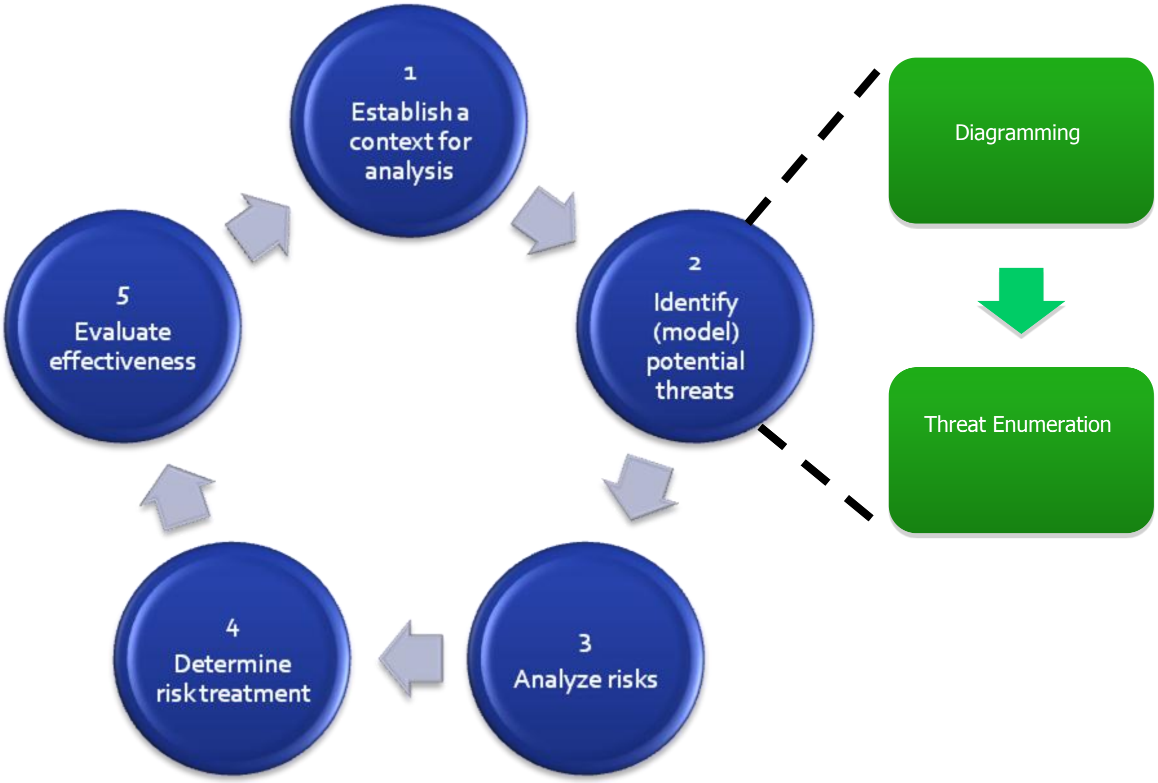


Figure 4. The risk/gap analysis process.

Organizations should also identify the stage or stages of the information lifecycle that correspond to each flow of data between two systems, or between a system and a storage unit. For highly complex cases, it might be necessary to draw multiple DFDs, each focusing on one stage or a few stages of the information lifecycle.

Threat enumeration is the result of a systematic analysis of the DFD. In this context, a threat is not limited to attackers or technical threats; it can be anything that might violate the four data privacy and confidentiality principles.

[18] http://technet.microsoft.com/en-us/library/dd941826.aspx

Step 3: Analyzing risk. Most organizations have already taken steps to ensure data security and privacy as specified by their existing control framework and/or ISMS. To complete this step, an organization should first gather information about its own and the CSP's protective technologies and activities. Then, for each cell in the Risk/Gap Analysis Matrix, it should determine which technologies and activities support compliance with each of the four privacy and confidentiality principles. This step concludes when threats that are not addressed by existing protective measures have been identified and the related risks have been evaluated.

Step 4: Identifying mitigation measures. In the appropriate cells in the matrix, the organization should list additional technologies and activities that are necessary to bring each risk to an acceptable level and do a cost/benefit evaluation. The step ends with a decision on whether and how each identified risk will be mitigated, transferred, or assumed.

Step 5: Evaluating the effectiveness of mitigation measures. This step involves reviewing the results of the preceding steps and reinitiating the cycle if unacceptable risks remain.[19]

A Final Note on Accountability and Risk Management

While the CSP should be responsible for living up to the promises it has made and for meeting the requirements specified in the service-level agreement, the organization that hires a CSP should hold itself accountable for maintaining the security and privacy of confidential data. Laws and regulations typically define the collecting organization as the responsible party, and as recent history shows, so does public opinion. Looking at one's own organization as the accountable party for data security and privacy is, quite simply, sound risk management.

Conclusion

Moving systems and processes into the cloud involves a tradeoff between benefits (such as flexibility, scalability, potential improvement of key aspects of data security, and a more desirable cost structure) and potential challenges (such as diminished control over data and over the infrastructure that houses and processes that data). Process transparency on the part of the CSP will not only help organizations understand how the CSP will protect data security and privacy, but it can also help them meet regulatory compliance obligations and internal policy requirements.

Organizations that want to move confidential data to the cloud should systematically identify risks to data privacy and security in the information lifecycle and risks of noncompliance:

- Before moving to the cloud, organizations should understand the legal implications of moving certain types of data to a specific CSP in a specific geographic location. CSP rights and potential ownership of data should be clarified and reviewed with legal counsel.

[19] For a more detailed description of the risk/gap analysis process, see "A Guide to Data Governance for Privacy, Confidentiality, and Compliance: Part 3: Managing Technological Risk" at www.microsoft.com/datagovernance.

- Organizations should confirm the long-term viability of the CSP, along with the implications and costs of potentially switching to a new CSP in the future.
- Organizations should ask CSPs to substantiate their claims about the level and quality of security, privacy, and compliance mechanisms that they can provide. CSPs can provide a reasonable level of transparency in the form of documentation—backed by relevant third-party certifications—about protective measures and processes and how their effectiveness and execution will be verified.

Organizations should also diagram their data flows to better understand the security, privacy, and compliance risks in the context of the cloud. To do this, they should use the Risk/Gap Analysis Matrix to identify gaps in their existing protective and compliance measures and those of the CSP. The matrix helps identify threats and residual risks in specific data flows—valuable information that can be used to improve protections and manage associated risks.

Glossary of Terms

accountability In the context of privacy, the principle that an organization should be responsible for complying with measures that give effect to its privacy principles and policies. In the context of data security, this principle refers to the implementation of controls, technologies, and processes that enable the organization to make privacy and security transgressors accountable for their actions.

attestation A statement or opinion by an auditor as to whether an assertion is true.

authority document Any document containing control requirements applicable to an organization, including but not limited to governance, standards, and contractual requirements.

data governance The exercise of authority, control, and shared decision making (planning, monitoring, and enforcement) over the management of data assets.[20]

data protection The management of personal information. In the United States, *privacy* is the term used in policies, laws, and regulations. In the European Union and other countries, the term *data protection* is more often used in privacy-related laws and regulations.[21]

elasticity The ability of cloud services to rapidly increase and decrease the amount of computer resources as needed, as well as the ability to release them when they are no longer needed.

GRC Governance, risk management, and compliance.

- **Governance** ensures that the business focuses on core activities, clarifies who in the organization has the authority to make decisions, determines accountability for actions and responsibility for outcomes, and addresses how expected performance will be evaluated. All of this happens within a clearly defined context that might span a division, the entire organization, or a specific set of cross-discipline functions.
- **Risk management** is a systematic process for identifying, analyzing, evaluating, remedying, and monitoring risk. As a result of this process, an organization or group might decide to mitigate a risk, transfer it to another party, or assume the risk along with its potential consequences.
- **Compliance** generally refers to actions that ensure behavior that complies with established rules as well as the provision of tools to verify that behavior. It encompasses laws as well the organization's own policies, which in turn can be based on best practices. Compliance requirements are not static, and compliance efforts should not be either.

Infrastructure as a Service (IaaS) A type of cloud service in which a CSP provides access to basic IT infrastructure (network, hardware, core operating system, and virtualization software) on which the organization can deploy its own applications and data in a virtualized environment. Examples include Amazon's EC2 and Rackspace's Cloud Servers.

[20] *The DAMA Dictionary of Data Management*, 1st Edition, 2008.

[21] IAPP Information Privacy Certification: Glossary of Common Privacy Terminology, International Association of Privacy Professionals (IAPP), 2006.

multi-tenancy The use by multiple customers of cloud services on a single server, typically through the use of virtualization software.

personal data Any and all data that relates to an identifiable individual.[22]

personal information Any information that 1) relates to an individual and 2) identifies or can be used to identify the individual. Such information might include an individual's name, postal address, e-mail address, telephone number, Social Security number, or other unique identifier.

personally identifiable information (PII) Any information that can be traced to a particular individual. Usually this type of information is identified through an identification block of data, such as a name, mailing address, phone number, Social Security number, or e-mail address. Personal user preferences tracked by a Web site via a cookie are also considered personally identifiable when linked to other PII provided by a user online.

Platform as a Service (PaaS) A type of cloud service in which a CSP provides the underlying infrastructure, including operating systems and storage, and allows organizations to create or run applications using languages and tools supported by the CSP. An example is Microsoft's Windows Azure platform.[23]

privacy The appropriate use of personal information under the circumstances. What is appropriate will depend on the context, laws, and the individual's expectations. Privacy also refers to the right of an individual to control the collection, use, and disclosure of personal information.

risk management Managing a situation or project so that minimum loss or damage will result if the risk materializes.[24]

sensitive personal information/sensitive data The 1998 EU Directive distinguishes between ordinary personal data, such as name, address, and telephone number, and sensitive personal data, such as racial or ethnic origin, political opinions, religious beliefs, trade union membership, health, sex life, and criminal convictions. Under the act, the processing of sensitive data is subject to stricter conditions.[25]

Software as a Service (SaaS) A type of cloud service in which a CSP provides and manages software applications. An example is Microsoft Online Services, which include hosted versions of Microsoft Exchange and Microsoft SharePoint for use by both public and private organizations.

threat modeling As used in this series of whitepapers, a technique for determining and analyzing threats to security, privacy, and compliance.

virtualization Simulation of underlying computer hardware that allows software, typically one or more guest operating systems, to run unmodified.

[22] Ibid.

[23] www.microsoft.com/windowsazure

[24] *The DAMA Dictionary of Data Management*, 1st Edition, 2008.

[25] IAPP Information Privacy Certification: Glossary of Common Privacy Terminology, International Association of Privacy Professionals (IAPP), 2006.

Appendix: Example Risk/Gap Analysis Matrix

Scenario: An Internet merchant plans to launch a new Web site that will sell inexpensive jewelry directly to consumers. As part of its strategic plan, the company wants to move the new system to the cloud and has begun evaluating CSPs.

One of the CSPs has provided documentation with an overview of its ISMS and an explanation of how it approaches compliance obligations.[26] The CSP has also agreed to provide, under a nondisclosure agreement, copies of its third-party-audited ISO/IEC 27000 compliance certification and SAS 70 attestations, once a service agreement is signed.

After reviewing the documentation and the CSP's financial statements (which are publicly available because the CSP is a shareholder-owned company listed on the stock exchange), the merchant decides to hire that CSP.

The CSP's documentation shows that privacy and security requirements are considered in new applications from design to deployment. The merchant wants to take advantage of this platform, but because of time pressures it decides to initially launch the site with a mixed configuration. The Web front end, along with all code related to capturing consumer data, will run on a CSP-managed server in the cloud, while the store application itself will run on the merchant's on-premises servers. All customer PII and transactional data will be stored in the cloud, as shown in Figures 5 and 6 on the next page. Note that for security reasons, a log server will be placed in a separate physical location from that of the application server. The log server will keep a time-stamped record of every transaction, including date, time, type of transaction, and total cost, but no customer PII or transaction details.

In the future, once the new version of the application is ready, the merchant will complete the move to the cloud.

The merchant's IT staff applies the risk/gap analysis process described earlier in this paper, identifying threats and risks and selecting mitigation measures. The identified threats are shown in Table 1 on the following pages; the final version of their Risk/Gap Analysis Matrix is shown in Table 2. (Note that these tables refer to users of the site as "customers.")

As you review this scenario, keep in mind that some threats can be classified under one or more of the privacy and confidentiality principles or within multiple lifecycle stages. Similarly, risk mitigation measures can be classified under multiple lifecycle stages. To keep the example simple, we have opted to not repeat the same threat and/or mitigation measure multiple times.

[26] For real-life examples of what this documentation looks like, see "Microsoft's Compliance Framework for Online Services" and "Securing Microsoft's Cloud Infrastructure" at www.globalfoundationservices.com/security/index.html. See also www.microsoft.com/downloads/details.aspx?FamilyID=5736aaac-994c-4410-b7ce-bdea505a3413&displaylang=en.

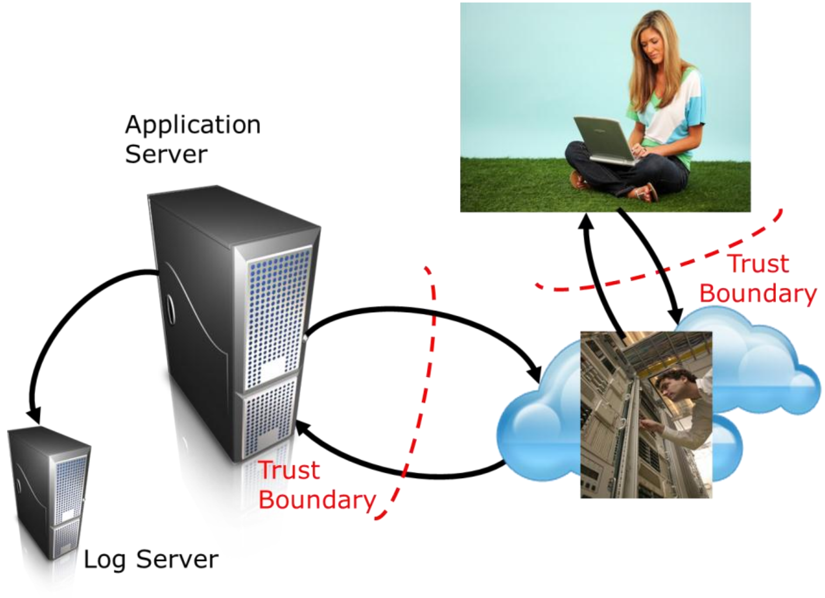


Figure 5. The Collection and Update stages in the risk/gap analysis process, with corresponding data flows and trust boundaries. Note that there is no trust boundary between the application server and the log server because they both belong to the merchant and operate under the same jurisdiction (set of policies, processes, and procedures) and are subject to the same set of laws and regulations.

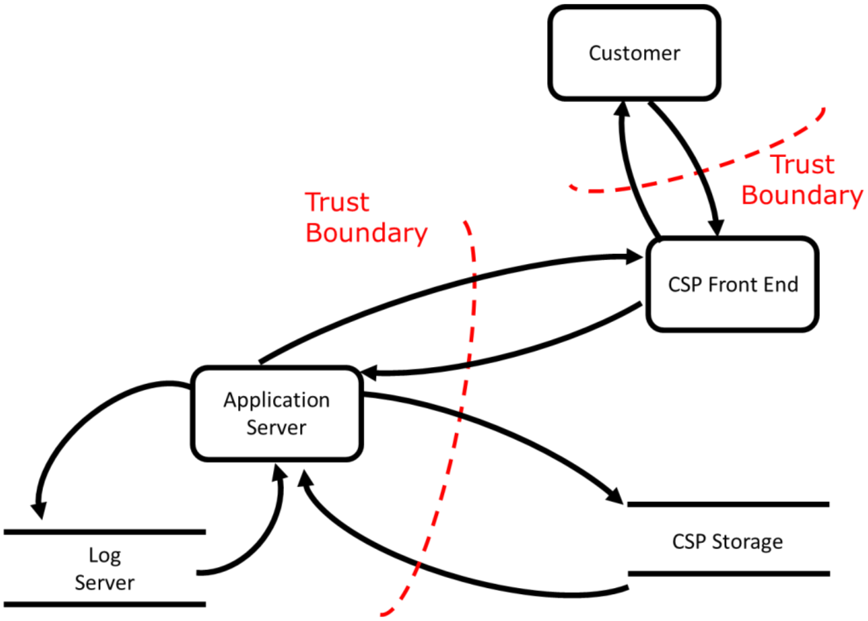


Figure 6. This diagram represents the data flow shown in Figure 5. Rectangles denote entities (computer systems in this case), parallel bars denote storage systems, and arrows indicate data flows.

Table 1. Threats Identified by the Merchant

Lifecycle Stage(s)	Data Privacy and Confidentiality Principle	Threat Type	Threat Description
Collect/Update	1	Choice and Consent	Customers might not be notified of the merchant privacy policy in a timely manner and/or that policy might not be consistent with the CSP's privacy policy.
Collect/Update	1	Choice and Consent	Customer choice and consent options might not be clearly described.
Collect/Update	1	Accountability	Customer PII and transactional information might not be classified appropriately.
Store	1	Accountability	Ownership of metadata and aggregated data derived from the merchant's customer data is not clear.
Store	3	Compliance	PII and/or transactional data might be lost or corrupted while in storage.
Collect/Update	1	Compliance	The data classification policy does not consider storage of PII.
Collect/Update	2	Information Protection	PII and transactional data will be transferred from the customer to the CSP and then to the application server. Data might be intercepted by a third party.
Store	3	Information Protection	Customer PII might be breached while being stored by the CSP.
Collect/Update	2	Compliance	Communication with the log server might be lost during operation.
Collect/Update	2	Data Quality	Customers might not be offered a way to update their choice and consent options.
Collect/Update	2	Data Quality	Data might be corrupted or made inaccessible by CSP software patching and/or hacker attacks.
Collect/Update	3	Accountability	No data breach notification process is in place with the CSP.
Collect/Update	3	Accountability	No incident response plan is in place with the CSP in case of internal/external attack.
Collect/Update	3	Compliance	Data classification and privacy policies are not updated and documented to reflect the new CSP scenario.
Process	2	Information Protection	Customer PII might be used by marketing for purposes not authorized by the customer.
Store	4	Compliance	Compliance reports, the reporting schedule, and report recipients within organization are not defined.
Store	4	Compliance	The merchant has no internal escalation path for unforeseen incidents.
Delete	4	Information Protection	The merchant has no log data retention policy.

Table 2. The Risk/Gap Analysis Matrix

Information Lifecycle Stage and Description		Secure Infrastructure	Identity and Access Management	Information Protection	Auditing and Monitoring	Manual Controls
Collect/ Update	Information is collected directly from the customer, who provides personal data and selects purchases. Order changes and cancellations are managed using the same process.	Application and log servers are on regular operating system and application patch cycles and are up-to-date on anti-malware signatures. **CSP-based front end links to the merchant's privacy policy on the application server.** **CSP-based front end explains customer choice and consent options per the merchant's standards and allows customers to update their PII.** **Incoming data is classified and tagged by the application per the customer's choice and consent and according to the new data classification policy.**	User creates account and password. Access privileges are administered according to the merchant's policy (role-based, least privilege) on on-premises Active Directory® servers. Account provisioning and administration are managed in-house per the merchant's policy. Access to CSP-stored data is controlled through federation, as specified in the SLA. **The list of individuals with access privileges is reviewed every 90 days.**	Transaction log data is encrypted in transit and at rest. **Communications involving PII (between the customer and the CSP and between the CSP and the merchant) are encrypted using SSL.** **User choice and consent data is stored in the cloud, along with PII and transactional data.**	All material transactions are logged according to the merchant's logging framework. Communications channel to log servers and log server activity are monitored. A failover process to local log servers in processor facilities is up and running. **Data classification and privacy policies are updated to include PII stored by the CSP.**	Verify that the CSP's privacy policy is aligned with that of the merchant. A process and schedule for periodic verification of individual and group access rights are in place and followed.

Information Lifecycle Stage and Description		Secure Infrastructure	Identity and Access Management	Information Protection	Auditing and Monitoring	Manual Controls
Process	Processing is limited to the extraction of the minimum data needed for marketing purposes. No third parties will have access to customer data.	Servers are on regular operating system and application patch cycles and are up-to-date on anti-malware signatures.	Access to the marketing application is limited to a pre-specified group, and privileges are granted per the merchant's policy (role-based, least privilege).	Data is directly accessed by the application over an encrypted communications channel, following parameters provided by the marketing team.	**All executions of the marketing application, including the identity of the requesting party, are logged according to the merchant's logging framework.**	
Delete				**Customer requests for account deletion are stored in log servers and deleted from the cloud database per the SLA.**	Log data is deleted per the data retention policy.	**Accepted data deletion methods are specified in the SLA and verified through certification.**
Store	All transactional details and customer information are stored with the CSP. Storage requirements for the CSP are specified in the Transfer phase.	Log servers are on regular operating system and application patch cycles and are up-to-date on malware signatures. CSP infrastructure protection (including firewalls, anti-malware software, and patching) takes place according to CSP processes and procedures specified in the CSP's ISMS and has ISO/IEC 27000 certification and SAS 70 Type I attestation.	Access to transaction logs and transaction log reports is granted on a per-role basis. Log access privilege lists are reviewed periodically, per the organization's policy.	Transaction log data is encrypted while at rest. All transactions with the CSP take place over an encrypted communications channel. SQL Server row-level encryption is used in all cases. Data is encrypted before transmission. Data backup process and integrity requirements, along with information destruction requirements, are specified contractually following standards specified by the CSP's ISMS.	Log failover data transfer and backup procedures are in place. The log backup process and schedule are in place. Access and use reports from the CSP are defined, and a schedule is in place.	Clarify ownership of aggregated data and metadata with the CSP. Reflect conclusion in the terms of agreement. CSP incident response and notification processes are in place, as specified by the SLA and the CSP's ISMS. Compliance requirements are verified through third-party certifications and attestations, along with reports specified in the SLA. A regular schedule for reviewing CSP access logs is defined and followed.

References

Application Privacy Assessment: www.microsoft.com/datagovernance

"A User-Centric Identity Metasystem," Cameron, Posch, and Rannenberg. October 5, 2008, www.identityblog.com/wp-content/images/2009/06/UserCentricIdentityMetasystem.pdf

Cloud security:

- Cloud Security Alliance (CSA): www.cloudsecurityalliance.org
- European Network and Security Agency (ENISA): www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment
- Cloud Computing Security Considerations: http://go.microsoft.com/?linkid=9708479

Compliance Solution Accelerators: http://technet.microsoft.com/en-us/solutionaccelerators/dd229342.aspx

***DAMA Dictionary of Data Management*, 1st Edition, USA, 2008,** www.dama.org/i4a/pages/index.cfm?pageid=3345

DataLossDB: datalossdb.org

Generally Accepted Privacy Principles, American Institute of Certified Public Accountants (AICPA) and Canadian Institute of Chartered Accountants (CICA): www.aicpa.org/InterestAreas/InformationTechnology/Resources/DataIntegration/DownloadableDocuments/GAPP_%20Practitioner_%20092006.pdf

IAPP Information Privacy Certification: Glossary of Common Privacy Terminology, International Association of Privacy Professionals (IAPP), 2009, https://www.privacyassociation.org/certification/free_resources

"Information Accountability," Weitzner, Abelson, Berners-Lee, Feigenbaum, Hendler, and Sussman, Massachusetts Institute of Technology CSAIL Technical Report, June 2007, http://dspace.mit.edu/bitstream/handle/1721.1/37600/MIT-CSAIL-TR-2007-034.pdf?sequence=2

Microsoft Online Services: www.microsoft.com/online/products.mspx

Microsoft Privacy Guidelines for Developing Software Products and Services: http://download.microsoft.com/download/0/8/2/082448D8-2AED-45BC-A9A0-094840E9E3A2/Microsoft_and%20Privacy_guidelines_for_developers.doc

"Microsoft's Compliance Framework for Online Services," www.globalfoundationservices.com/security/index.html

Microsoft Security Development Lifecycle: www.microsoft.com/security/sdl/default.aspx

Microsoft Windows Azure Platform: www.microsoft.com/windowsazure

NIST Definition of Cloud Computing: http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc

OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, Organisation for Economic Co-operation and Development, www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html

"Privacy in the Clouds: Risks to Privacy and Confidentiality from Cloud Computing," R. Gellman, World Privacy Forum, www.worldprivacyforum.org/cloudprivacy.html

Risk management:

- Information Risk Analysis Methodology (IRAM), https://www.securityforum.org/services/publictools/publiciram
- Risk Management Guide for Information Technology Systems, National Institute of Standards and Technology, U.S. Department of Commerce, http://csrc.nist.gov/publications/nistpubs/800-30/sp800-30.pdf
- Standard AS/NZS 4360:2004, http://infostore.saiglobal.com/store/Details.aspx?docn=AS0733759041AT

"Securing Microsoft's Cloud Infrastructure," www.globalfoundationservices.com/security/index.html

"Security Features in Microsoft Online Services," www.microsoft.com/downloads/details.aspx?FamilyID=5736aaac-994c-4410-b7ce-bdea505a3413&displaylang=en

Threat modeling:

- "Experiences Threat Modeling at Microsoft," M. Shostack, Microsoft, 2008, www.homeport.org/~adam/modsec08/Shostack-ModSec08-Experiences-Threat-Modeling-At-Microsoft.pdf
- Microsoft's IT Infrastructure Threat Modeling Guide, http://technet.microsoft.com/en-us/library/dd941826.aspx