NATIONAL ARCHIVES

NARA Bulletin 2010-05

September 08, 2010

TO: Heads of Federal agencies

SUBJECT: Guidance on Managing Records in Cloud Computing Environments

EXPIRATION DATE: Expires when revoked or superseded

1. What is the purpose of this bulletin?

This bulletin addresses records management considerations in cloud computing environments and is a formal articulation of NARA's view of agencies' records management responsibilities. As agencies are increasingly evaluating, piloting, and adopting these technologies, they must comply with all Federal records management laws, regulations, and policies.

2. How does this bulletin differ from "Frequently Asked Questions about Managing Federal Records in Cloud Computing Environments"?

NARA issued an FAQ in February 2010 to provide agencies with a basic overview of cloud computing. This bulletin expands on that discussion by including a more detailed definition, Federal agency examples of cloud computing, records management guidelines, and contract language to consider when procuring cloud computing services.

3. What is cloud computing?

Cloud computing is a technology that allows users to access and use shared data and computing services via the Internet or a Virtual Private Network. It gives users access to resources without having to build infrastructure to support these resources within their own environments or networks.

General interpretations of cloud computing include "renting" storage space on another organization's servers or hosting a suite of services. Other interpretations of cloud computing reference particular social media applications, cloud-based e-mail, and other types of Web applications. However, the National Institute of Standards and Technology (NIST) has been designated to develop standards and guidelines for the Federal cloud computing effort and to provide an authoritative definition.

NIST defines cloud computing as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal

management effort or service provider interaction." (NIST Definition of Cloud Computing, Version 15, 10-07-2009) NIST has stated that the definition of Cloud Computing is evolving. The user should consult the most current definition available from NIST and other resources.

NIST also identifies five essential characteristics of cloud computing:

- *On-demand self-service.* A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.

- *Broad network access.* Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs).

- *Resource pooling.* The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.

- *Rapid elasticity.* Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.

- *Measured Service.* Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported providing transparency for both the provider and consumer of the utilized service.

  (NIST Definition of Cloud Computing, Version 15, 10-07-2009)

The terminology above is used in the IT community and by NIST to describe characteristics of cloud computing.

4. What are cloud computing service and deployment models?

Cloud computing service models refer to how an agency can adopt cloud computing. Currently NIST describes the models as follows:

- *Cloud Software as a Service (SaaS).* The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure

including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

- *Cloud Platform as a Service (PaaS).* The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

- *Cloud Infrastructure as a Service (IaaS).* The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

  Depending upon user needs, and other considerations, cloud computing services are typically deployed using one of the following four models as defined by NIST:

- *Private cloud.* The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise.

- *Community cloud.* The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise.

- *Public cloud.* The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

- *Hybrid cloud.* The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).

(All definitions are from NIST Definition of Cloud Computing, Version 15, 10-07-2009)

Public and private clouds are terms used in the IT community and by NIST to describe various cloud configurations. These terms are not related to whether records in those clouds are publicly accessible. In addition, definitions do not preclude agency responsibilities to manage the records.

5. How are Federal agencies using cloud computing?

NARA interviewed agencies that are using cloud computing services to achieve benefits such as cost savings, accessibility, scalability, collaboration, and flexibility. All of the agencies interviewed recognize the challenge of managing data in the cloud. They have begun to address concerns around cloud computing environments such as privacy, security, and data ownership. However, the agencies stated that managing records in a cloud computing environment is a concern that they are only beginning to address.

In one example, an agency dealing with globally-dispersed employees needed a rapid solution for information sharing. Using a commercial contractor, the agency deployed a private cloud to share financial data, capture reports, provide world-wide access to information, and solve security challenges. In this instance, the agency used a separate commercial platform to prevent unauthorized users from using the application as a back door to access secure agency servers. The agency said the system meets all fiscal audit requirements, but also said it recognizes that records management requirements were not addressed, and no data is being deleted.

In another example, at least two units in the same agency built and offer cloud computing services as providers to other offices in different parts of the agency. Both units deal with classified information and need to retain control of these records within their organizations. They control the use of and access to the system. One of the units experienced an increase in the use of its collaboration software because customers no longer need to worry about software and hardware acquisition, updating software and operating systems, back-ups, and access/permissions control. Customers are responsible for the content and determining its record status and for managing records in the cloud.

In a final example, at least two other agencies offer cloud computing services both within the agency and to other agencies. One of these agencies is offering storage space over which the customer has complete configuration control. While still a pilot, it is anticipated that customers will be responsible for managing records stored in the cloud. The two agencies are considering "Terms of Service" agreements that would address records management requirements.

For more agency examples see the CIO Council's report on the State of Public Sector Cloud Computing (pdf). For security issues with cloud computing see GAO's report Federal Guidance Needed to Address Control Issues with Implementing Cloud Computing (pdf).

6. What are some of the records management challenges associated with cloud computing?

NARA recognizes that the service and deployment models affect how records may be created, used, and stored in cloud computing environments. The level of agency control over the design,

implementation, and operations of a cloud computing environment containing Federal records may vary depending on service and deployment models. For example, in the case of IaaS and PaaS service models, the agency is more likely to maintain the records outside the cloud. In the SaaS service model, the agency may maintain records in contractor-provided clouds, and any negotiated contract language will need to address specific records management responsibilities. See Question 8 for a general clause that an agency can modify to fit the planned type of service.

NARA has identified several records management challenges with cloud computing environments:

- Cloud applications may lack the capability to implement records disposition schedules, including the ability to transfer and permanently delete records or perform other records management functions. Therefore specific service and deployment models may not meet all of the records management requirements of 36 CFR Part 1236 (formerly 36 CFR part 1234). Examples of these requirements include:
  - Maintaining records in a way that maintains their functionality and integrity throughout the records' full lifecycle
  - Maintaining links between the records and their metadata
  - Transfer of archival records to NARA or deletion of temporary records according to NARA-approved retention schedules.

- Depending on the application, cloud service providers must be made aware of the record retention requirements governing a given body of Federal records stored in one or more cloud locations. Agencies need to be able to control any proposed deletion of records pursuant to existing authorities, wherever the records may be located in the providers' cloud. Cloud service providers must also act to ensure that records are accessible so as to ensure agency responsiveness to discovery, or FOIA/Privacy Act, or other access requests.

- Various cloud architectures lack formal technical standards governing how data are stored and manipulated in cloud environments. This threatens the long-term trustworthiness and sustainability of the data.

- A lack of portability standards may result in difficulty removing records for recordkeeping requirements or complicate the transition to another environment. This could affect the ability of agencies to meet their recordkeeping responsibilities for temporary or historically valuable records being transferred to NARA.

- Agencies and cloud service providers should anticipate how continued preservation and access issues will be resolved in a contingency where the cloud service provider's business operations materially change (e.g., bankruptcy), or cease altogether.

7. How can agencies meet their records management responsibilities?

Federal agencies are responsible for managing their records in accordance with NARA statues including the Federal Records Act (44 U.S.C. Chapters 21, 29, 31,33) and NARA regulations (36 CFR Chapter XII Subchapter B). This is true regardless of which cloud service and deployment models are adopted. However, NARA recognizes that the differences between models affect how and by whom (agency/contractor) records management activities can be performed.

The following are guidelines for creating standards and policies for managing an agency's records created, used, or stored in cloud computing environments:

a. Include the agency records management officer and/or staff in the planning, development, deployment, and use of cloud computing solutions.

b. Define which copy of records will be declared as the agency's record copy and manage these in accordance with 36 CFR Part 1222. Remember, the value of records in the cloud may be greater than the value of any other set because of indexing or other reasons. In such instances, this added value may require designation of the copies as records.

c. Include instructions for determining if Federal records in a cloud environment are covered under an existing records retention schedule.

d. Include instructions on how all records will be captured, managed, retained, made available to authorized users, and retention periods applied.

e. Include instructions on conducting a records analysis, developing and submitting records retention schedules to NARA for unscheduled records in a cloud environment, These instructions should include scheduling system documentation, metadata, and related records.

f. Include instructions to periodically test transfers of Federal records to other environments, including agency servers, to ensure the records remain portable.

g. Include instructions on how data will be migrated to new formats, operating systems, etc., so that records are readable throughout their entire life cycles. Include in your migration planning provisions for transferring permanent records in the cloud to NARA. An agency choosing to pre-accession its permanent electronic records to NARA is no longer responsible for migration except to meet its business purposes.

h. Resolve portability and accessibility issues through good records management policies and other data governance practices. Data governance typically addresses interoperability of computing systems, portability of data (able to move from one system to another), and information security and access. However, such policies by themselves will not address an agency's compliance with the Federal Records Act and NARA regulations.

8. What is an agency's responsibility when dealing with contractors?

Ultimately, an agency maintains responsibility for managing its records whether they reside in a contracted environment or under agency physical custody (see 36 CFR Part 1222.32 (b)). When dealing with a contractor, an agency must include a records management clause in any contract or similar agreement. At a minimum, a records management clause ensures that the Federal agency and the contractor are aware of their statutory records management responsibilities.

The following is a general clause that an agency can modify to fit the planned type of service and specific agency records management needs:

Use of contractor's site and services may require management of Federal records. If the contractor holds Federal records, the contractor must manage Federal records in accordance with all applicable records management laws and regulations, including but not limited to the Federal Records Act (44 U.S.C. chs. 21, 29, 31, 33), and regulations of the National Archives and Records Administration (NARA) at 36 CFR Chapter XII Subchapter B). Managing the records includes, but is not limited to, secure storage, retrievability, and proper disposition of all federal records including transfer of permanently valuable records to NARA in a format and manner acceptable to NARA at the time of transfer. The agency also remains responsible under the laws and regulations cited above for ensuring that applicable records management laws and regulations are complied with through the life and termination of the contract.

If an agency decides to create or join a private or community cloud, it will still need to meet records management responsibilities. The agencies may describe these responsibilities in agreements among the participating offices or agencies. If a cloud provider ceases to provide services an agency must continue to meet its records management obligations. Agencies should plan for this contingency.

9. Where do I go for more information?

NARA's National Records Management Program can provide assistance. See "List of NARA Contacts for Your Agency." In addition, NARA maintains the Toolkit for Managing Electronic Records as a resource for agencies to share and access records management best practices.

DAVID S. FERRIERO
Archivist of the United States