



**Con los pies en la tierra:  
Un enfoque práctico a la nube**  
Nueve factores a considerar al evaluar  
el almacenamiento en la nube

por seth anderson  
consultor  
avpreserve  
2014



## Introducción

La planificación y toma de decisiones en cualquier proyecto de archivo y de preservación involucra considerar y reconsiderar los detalles. En la tarea de conservar y difundir colecciones valiosas en un entorno en constante cambio, a menudo es abrumador determinar las herramientas que deben utilizarse para estas responsabilidades. Los archivos y archivistas podrían identificar las soluciones apropiadas para la preservación y el acceso solo definiendo de los requisitos de la organización y la colección con el fin de llevar a cabo un cuidadoso escrutinio de los servicios disponibles. Esto es especialmente cierto cuando se consideran las tecnologías emergentes, que varían ampliamente en sus funciones y que representan nuevas exigencias y riesgos potenciales. En el caso de los servicios de terceros, tales como el almacenamiento en la nube, la importancia del proceso de elección es evidente. Asegurarse que un servicio puede realizar las funciones necesarias, y de manera, es el primer paso para comprender el ahorro de recursos y las ventajas que estos servicios proporcionan.

Al evaluar a los proveedores de almacenamiento en la nube, es peligroso asumir que tales servicios son solo de almacenamiento y, por lo tanto, sin complicaciones, o que las necesidades de almacenamiento son obvias y, por lo tanto, que son intrínsecamente satisfechas por el proveedor de servicios. La experiencia en cualquier selección de tecnología demostrará lo contrario. No hay dos servicios idénticos, y la variación entre los servicios a menudo representa la diferencia entre una experiencia exitosa y otra fallida. Nunca compre un servicio sin un escrutinio adecuado; las decisiones sin la información adecuada representan un riesgo de pérdida de tiempo, dinero e incluso activos.

No existe una solución todo-en-uno que satisfaga las necesidades de almacenamiento de preservación de todos los archivos. Con frecuencia, los servicios de almacenamiento en la nube desempeñan una parte de la infraestructura general de preservación de una organización, proporcionando un respaldo seguro para copias de preservación, o apoyando la disponibilidad de archivos de acceso de almacenamiento de baja latencia. Por lo tanto, el escrutinio y la selección deberán estar alineadas con de las necesidades organizacionales y de recolección con el ofrecimiento y la funcionalidad de un servicio. Esto significa definir sus criterios de aceptación para una funcionalidad óptima, así como entender cómo un servicio se ajustará a su entorno de preservación.

La dificultad está en saber por dónde empezar y qué preguntas hacer. Los nueve temas siguientes proporcionan un punto de partida. Use estos temas para promover la investigación y el debate, así como para asegurarse de que las consideraciones relevantes a continuación sean debidamente documentadas en el contrato de servicio.

### **1. Definición de los requisitos**

El almacenamiento en la nube puede ser adecuado para satisfacer uno o más aspectos de un entorno de preservación digital integrado que consiste de sistemas, almacenamiento, políticas y personas. Comience por decidir a qué se ajusta el almacenamiento de terceros y dónde lo hace dentro de su entorno de preservación.

Entonces, es necesario entender exactamente qué papel desempeñará un servicio de terceros, los aspectos que el servicio no proporcionará (p. ej., validación de SIP, validación de la integridad de los datos, caracterización de archivos, entre otros) y cómo se puede integrar con otras tecnologías antes de investigar y evaluar posibles soluciones. Esto disminuye el número de posibles servicios a explorar y le asegura que no pierda su tiempo explorando los servicios que no pueden satisfacer sus necesidades.

Con frecuencia, las organizaciones encuentran valor en el diseño de una arquitectura de almacenamiento de varios niveles y distribuida para servir a una amplia gama de cuestiones que abarcan consideraciones financieras, necesidades de los usuarios, redundancia, separación geográfica y otros requisitos funcionales y comerciales. Hay muchas combinaciones posibles de arquitecturas de almacenamiento de varios niveles, y diversos factores impulsarán las decisiones respecto de la mejor opción para una organización determinada. La mejor práctica típica en los archivos es mantener diferentes atributos de acervos para satisfacer los requisitos de preservación y acceso. Estos tienen diversos usos previstos, cuestiones de seguridad y

requisitos de desempeño, todos los cuales dictan las necesidades de almacenamiento y de ancho de banda. Por ejemplo, una copia de acceso de video, destinado para la navegación por parte de un estudiante en una universidad, se satisface de mejor manera por medio del almacenamiento en línea de alta disponibilidad y baja latencia para proporcionar acceso acelerado. Este conjunto de requisitos puede ser satisfecho adecuadamente tanto por soluciones in-situ como por servicios en la nube de terceros. Seleccionar cuál funciona mejor dependerá de factores técnicos y financieros expresados más detalladamente, los cuales se identifican de mejor manera a través del escrutinio en base a los requisitos.

**Preguntas que se deben formular:**

- ¿Con qué frecuencia se accederá a los materiales y por qué tipos de usuarios?
- ¿Con qué rapidez deben recuperarse los materiales?
- ¿Hay especificaciones particulares que deben cumplirse por seguridad, confiabilidad y tiempo de disponibilidad?
- ¿Cuál es la necesidad de almacenamiento prevista para los próximos 3 a 5 años?
- ¿Existen políticas sobre el tipo de almacenamiento requerido (p. ej., cinta vs. tecnología de disco)?
- ¿Existen políticas sobre la ubicación geográfica de almacenamiento o los gastos del presupuesto (p. ej., todos los servidores o el dinero gastado deben estar dentro de un estado, región, país específicos)?
- ¿Cuál es el presupuesto inmediato, a mediano y a largo plazo para la arquitectura de almacenamiento?
- ¿La organización prefiere los gastos de capital sobre los pagos por servicios, o viceversa?
- ¿La organización dispone de personal especializado para administrar los servicios de almacenamiento y preservación?
- ¿Está utilizando otros servicios en la nube que harían ventajoso utilizar el almacenamiento en la nube?

**2. Evaluación del proveedor**

El número de proveedores en el mercado de servicios en la nube sigue creciendo a medida que la computación en la nube se vuelve más omnipresente. Crear una lista corta de proveedores de servicios puede ser difícil, incluso cuando se han identificado necesidades. Es necesario enfocarse inicialmente en la evaluación del proveedor, no del servicio. Comprender la viabilidad de una empresa, la madurez de su servicio y operaciones, y la satisfacción de sus clientes, le proporcionará una imagen de la idoneidad del servicio para sus necesidades.

Un buen indicador de la madurez y la viabilidad de un proveedor es la composición, el tamaño y la satisfacción de su base de usuarios. Los proveedores externos sirven a una serie de diferentes sectores, por lo que variará su conocimiento de las necesidades únicas de un entorno de preservación de material de archivo. No es necesario que entiendan sus necesidades a fondo, pero esto podrá mejorar su experiencia como cliente.

El tamaño de la base de clientes puede ser un factor positivo o negativo. Un servicio con una base de usuarios diversa y amplia puede ser evidencia de un producto o servicio al cliente de calidad, aunque un grupo de clientes tan grande puede afectar la eficacia del servicio al cliente. Determinar si un proveedor es adecuado para su organización puede ser complicado, pero el tener en cuenta la relación de trabajo potencial es un paso necesario en cualquier proceso de selección de tecnología. Esto preparará el terreno para establecer la confianza hacia el servicio seleccionado.

**Preguntas que deben hacerse:**

- ¿Por cuántos años ha estado proporcionando este servicio el proveedor?
- ¿La compañía será financieramente viable en cinco años? ¿En diez años?
- ¿Cuáles son los protocolos de servicio al cliente del proveedor? ¿Cuál es el tiempo de respuesta establecido para responder a los problemas de los clientes?
- ¿Cuán grande es su base de clientes? ¿A qué tipos de organizaciones y empresas sirve?
- ¿La terminología y el posicionamiento demuestran que entiende las responsabilidades y funciones requeridas de los archivos y la preservación?
- ¿Cómo ofrece apoyo la empresa? ¿Solo por correo electrónico? ¿Planes de apoyo para las compras?
- ¿Los términos y condiciones, o términos de servicio, son razonables y aceptables?

**3. Administración del hardware de almacenamiento**

Al profundizar en los aspectos prácticos de la evaluación de los servicios, el enfoque debería ser la adherencia del servicio (o flexibilidad) a los principios básicos de preservación digital. Entrar en los detalles del hardware y la tecnología de almacenamiento probablemente supera la experiencia de la mayoría de los archivistas, y se desvía de la consideración principal que nos ocupa: **¿El cuidador puede realizar las acciones de preservación y de acceso necesarias en colaboración con este proveedor?**

Con esto en mente, al evaluar el hardware de un proveedor, el énfasis no está exclusivamente en el tipo o la arquitectura, sino en la administración del hardware por parte del proveedor. En apoyo de la preservación, el proveedor debe:

- Replicar materiales en lugares geográficamente separados,
- Monitorear la salud y edad del hardware de almacenamiento (el ciclo de vida estimado de cintas y unidades es de 3 a 5 años), y
- Actualizar periódicamente el hardware de almacenamiento y migración de datos.

Estos son los principios fundamentales que se plantean en cualquier iniciativa de preservación digital, pero son doblemente importantes al evaluar las opciones de almacenamiento en la nube. Puede ser fácil asumir que la replicación y la separación geográfica son parte del servicio de un proveedor, pero esta creencia puede poner su colección en riesgo.

**Preguntas que deben hacerse:**

- ¿Los datos están respaldados y son redundantes a través de al menos dos centros de datos geográficamente separados? ¿Dónde están ubicados?
- ¿Cuáles son los tipos de almacenamiento que utiliza el proveedor (cinta, disco, combinación)? ¿Qué significa esto para las actualizaciones de hardware y la migración de datos?
- ¿Cuál es el protocolo de actualización de hardware del proveedor? ¿Con qué frecuencia actualiza el hardware (reemplaza las unidades de disco y/o cintas)?
- ¿Qué verificaciones y controles se emplean durante la migración de los datos? ¿Los resultados de la migración de ...hardware son... reportados o están disponibles para el cliente?
- ¿Los informes de salud del hardware/de los medios, con respecto de la ubicación de los datos almacenados del ...cliente, se ponen... a disposición del cliente?
- ¿Proporciona su servicio a través de un proveedor de almacenamiento existente, como Amazon, o el proveedor mismo posee y mantiene sus propios servidores?

- 4. Administración de datos** Al transferir algunos elementos de la preservación de datos a un tercero, puede haber algo de inquietud sobre lo que ocurre exactamente con los archivos una vez que entran en un nuevo entorno que está fuera de su control. El proceso de escrutinio y exploración debería eliminar la opacidad sobre la operación de un proveedor de servicios, siempre que sea posible. Debe mitigar cualquier inquietud respecto de que su colección se está colocando en una caja negra y está siendo manipulada de maneras desconocidas e imprevistas. Un servicio de terceros debe, como mínimo, poder devolver exactamente lo que usted aporta.

Una vez más, es clave apegarse a los principios básicos de preservación digital. Los proveedores pueden ser capaces de confirmar la integridad de sus archivos (verificación del estado actual del archivo, en comparación con una base de referencia conocida, como sumas de control) a la entrega y durante toda la administración continua del almacenamiento. Si es así, es importante la manera en que realizan esta tarea. Si no es así, se debe determinar cómo el cliente puede manejar la administración de la fijeza y la integridad del archivo, y lo que esto significa para sus operaciones y recursos.

**Preguntas que deben hacerse:**

- ¿Se realiza alguna modificación a los archivos al ingresar en el entorno de almacenamiento (p. ej., cifrado o deduplicación)?
- ¿El servicio garantiza la integridad de sus datos? Si es así, ¿cómo lo hace y con qué frecuencia? ¿Qué sucede si una comprobación de integridad de datos no supera la prueba?
- ¿Cómo el proveedor rastrea y los datos dentro de su sistema? ¿Generan metadatos que podrían usarse en las operaciones del lado del cliente? ¿Ofrecen informes al cliente?
- ¿Con qué estándares y/o certificaciones cumple el proveedor (p. ej., ISO 16363 - Auditoría y certificación de repositorios digitales)?

**5. Informes y metadatos**

La importancia de la documentación en la preservación digital no puede ser subestimada, pero los requisitos de informes y metadatos de una colección de archivo pueden ser desconocidos para la mayoría de los proveedores de almacenamiento externos. Los archivos deben tener la capacidad de dar cuenta de los eventos que ocurren en el ciclo de vida de un activo digital, rastrear los cambios en los datos (deseados o no) para garantizar la integridad de los materiales de la colección y solucionar cualquier problema que pueda surgir. La administración de estos factores se ve mitigada por el desarrollo de las políticas y los procedimientos de preservación aprobados por múltiples partes, que con frecuencia no son los autores de las políticas. Los informes del sistema pueden cerrar el ciclo, proporcionando al autor de la política y al administrador de la colección la evidencia de la aprobación y eficacia adecuadas de las políticas existentes. Este enfoque apoya la toma informada de decisiones para la administración futura de una colección y las políticas asociadas. Las características de los informes y metadatos del servicio que usted elija tendrán impacto en los recursos necesarios para proporcionar acceso continuo y la preservación, y no deben pasarse por alto.

**Preguntas que deben hacerse:**

- ¿Qué informes y registros están a disposición del cliente (p. ej., registros de errores, resultados de la verificación de fijeza, resultados de incorporación, informes de migración o registros de acceso)?
- ¿Qué metadatos se generan durante el almacenamiento (p. ej., verificaciones de fijeza, incorporación de detalles, informes de migración o detalles de replicación)?
- ¿Qué mecanismos (p. ej., informes enviados por correo electrónico, panel de administración, API) permiten a un cliente recuperar los informes y/o metadatos?

- 6. Desempeño** Una vez definida la arquitectura de la información, se debe determinar si un proveedor puede satisfacer sus parámetros de desempeño para el flujo hacia y desde el almacenamiento. Cada organización tendrá necesidades diferentes, pero los requisitos de los términos de desempeño generales serán determinados por la finalidad del servicio que usted está considerando. Si lo está utilizando para el acceso, sus requisitos probablemente necesitarán la entrega de materiales a alta velocidad desde el almacenamiento, así como la capacidad de soportar un número de usuarios con un período de latencia relativamente baja. Por el contrario, el almacenamiento de preservación quizás no necesite una alta disponibilidad de materiales, pero en los casos en que se requiere acceso como validación de la suma de comprobación o la creación de nuevos derivados, es esencial saber si el servicio coloca alguna restricción al acceso a los materiales en un almacenamiento profundo/de costo más bajo. Esto podría indicar si los escenarios de acceso previstos ocasionarán cargos adicionales para su organización (p. ej., cargos adicionales de Amazon Glacier para el acceso superior a **10TB** por mes, o precios que varían durante diversos períodos de recuperación). Algunos elementos de estas consideraciones estarán determinados por factores tecnológicos en su organización (p. ej., la arquitectura del sistema, las velocidades de conexión y de procesamiento, o el número de usuarios potenciales). Todos estos factores deben tomarse en cuenta para proporcionar una imagen completa de sus habilidades y necesidades de infraestructura existentes.

**Preguntas que deben hacerse:**

- ¿Cuál es el ancho de banda máximo permitido para flujos desde/hacia el servicio?
- ¿Cuál es el número estimado de usuarios finales que solicitarán acceso? ¿Número de empleados de administración que accederán al servicio? ¿Cuál es el número máximo de usuarios o conexiones que accederán al sistema simultáneamente?
- ¿A qué tipos de archivos accederá cada tipo de usuario (p. ej., imágenes fijas, documentos, audio, video) y qué tipo de acceso necesitan (p. ej., transmisión vs. descarga)?
- ¿El servicio proporciona la entrega de los bienes a los usuarios finales? Y, en caso afirmativo, ¿cómo lo hace (p. ej., transmisión, códigos integrados o la integración con los sistemas a través de la API)?
- ¿Cuáles son sus periodos de latencia aceptables para la recuperación y el acceso, y cuáles son los costos asociados?
- ¿Cuántos datos tiene autorizado a recuperar cada mes sin costo adicional? ¿Cuál es la escala móvil de tarifas una vez que se supera este límite?

- 7. Seguridad** El proveedor deberá garantizar la seguridad de los materiales para obtener grados importante de confianza en el cliente. Esto incluye garantizar que la información confidencial se maneje adecuadamente y esté a salvo de la intervención humana no deseada. Corresponde al cliente definir sus requisitos de seguridad y el nivel de recursos que está dispuesto a invertir en la administración continua de los protocolos de seguridad. Por ejemplo, algunos servicios pueden proporcionar características de seguridad, como el cifrado, pero es importante saber quién es el responsable de la administración de las claves de descifrado. Además de considerar la seguridad, esto requiere comprender el nivel de recursos disponibles para administrar y garantizar la seguridad de los materiales del lado del cliente. Al considerar las opciones de seguridad, es importante considerar también el impacto de los diversos mecanismos de seguridad en las estructuras de desempeño y permisos.

**Preguntas que deben hacerse:**

- ¿Los materiales en su colección contienen información sensible que requiere un aumento de la seguridad en un entorno de almacenamiento en la nube?
- ¿Qué protocolos de seguridad ha establecido el proveedor? ¿Quién es responsable de la administración y el mantenimiento de las claves?
- ¿Con qué estándares de seguridad cumplen y qué certificaciones han recibido (p. ej., ISO 27001 - Gestión de seguridad de la información, Estándares de seguridad de HIPAA para registros médicos, o el Programa Federal de Gestión de Riesgos y Autorizaciones [FedRAMP])?
- ¿El cliente puede solicitar y administrar sus propios protocolos de seguridad (p. ej., cifrado de clave)? ¿Cuáles son las implicaciones de los recursos por la administración de manera interna? ¿El cliente recibe apoyo para la implementación y administración de claves de cifrado y descifrado?

**8. Recuperación de desastres**

En estrecha relación con la confidencialidad está la seguridad de sus datos en casos de desastres, ya sean causados por el hombre o por fuerza mayor. Hay muchas historias de horror de fallas inesperadas en centros de datos, y es importante tener en cuenta las implicaciones de este tipo de eventos, pero es aún más valioso enfocarse en la respuesta del proveedor a los escenarios de desastres y evaluar métodos de notificación, el servicio y apoyo al cliente, y los planes de conciliación. En la mayoría de las situaciones el cambio a la copia de seguridad debe producirse sin problemas. El cliente debe entender cómo se ejecutan los procesos para garantizar que los materiales están respaldados y protegidos contra pérdidas, cómo ha utilizado y utilizará el proveedor los métodos de comunicación para informar de los problemas, y estar seguro de que se ha establecido una infraestructura para apoyar la prestación continua de los servicios como estaba previsto. Es importante que el cliente haga la distinción entre la protección de los datos y el tiempo de recuperación. Todas las organizaciones desearán tener la capacidad de recuperar todos los datos en el caso de un desastre, y tendrán requisitos diferentes en cuanto al tiempo que se necesita para recuperar sus datos después de un evento catastrófico. Estos tiempos de recuperación variables pueden venir con precios sumamente diferentes.

**Preguntas que deben hacerse:**

- ¿Cuáles son los mecanismos del servicio para la recuperación del almacenamiento de copia de seguridad en caso de un desastre (potencialmente, además de una copia replicada)?
- ¿Cuáles son los protocolos para hacer frente a la pérdida de datos y el retorno al servicio en casos de desastre? ¿Cuáles son las políticas y los detalles del acuerdo de servicio con respecto a la pérdida de datos?
- ¿Cuáles son los métodos de comunicación y las políticas del servicio en los casos de interrupción del servicio y/o la pérdida de datos?
- ¿Cuál es la cantidad máxima de tiempo entre que los datos se ingresan a la ubicación de almacenamiento primario y se replican a una ubicación de almacenamiento redundante geográficamente separada?
- ¿Cuál es el tiempo máximo que se necesita, desde el punto de desastre o pérdida del sistema de almacenamiento primario de datos, para la recuperación y el acceso a todos los datos presentes antes del desastre?

**9. Fin del servicio**

Explorar la extensión de un servicio incluye la revisión de todos los escenarios posibles en su ciclo de vida, incluyendo la rescisión por parte del cliente o el proveedor. Los acuerdos y protocolos de fin de servicio deben especificar la devolución de los materiales y los metadatos en su totalidad y tal como usted los entregó; una organización debe recibir lo que entregó (más los metadatos generados durante el almacenamiento). Este proceso puede ser generado por el término de servicio por parte del proveedor, en los casos en que el cliente no puede pagar o cuando el servicio ya no será ofrecido. El no identificar estas políticas puede presentar el riesgo de pérdida de materiales de la colección si no se comprende el final del servicio, ya sea a instancias del cliente o del proveedor. Además de identificar los protocolos, también es importante tener presente los costos.

**Preguntas que deben hacerse:**

- ¿Qué protocolos y mecanismos existen si el proveedor ya no ofrece el servicio?
- ¿Cómo se devuelven los datos al cliente o este los recupera al terminar el servicio?
- ¿Qué costos están relacionados con el retiro de los activos de este servicio?
- ¿Qué acciones llevan a la terminación del servicio y cómo es administrado esto por el proveedor del servicio?
- ¿El acuerdo de servicio documenta claramente esta información con suficiente detalle?

## Conclusión

Al considerar un servicio de terceros, tenga en cuenta los criterios señalados anteriormente pero no ignore los otros factores necesarios para la implementación de un entorno de preservación exitoso. Los servicios en la nube tienen un impacto limitado en su organización.

Un repositorio de preservación es una combinación de factores tecnológicos y humanos, procesos automáticos administrados y ejecutados mediante el manejo manual y las políticas. El servicio que seleccione será el que brinde a su personal las herramientas necesarias para apoyar de la mejor manera la preservación de su colección. El escrutinio es un proceso que requiere tiempo pero es valioso. Los recursos invertidos rendirán su fruto cuando el servicio en la nube se integre de manera exitosa en su entorno de preservación. Para ayudar en el proceso de toma de decisiones, AVPreserve publicará perfiles de los proveedores que resumen las capacidades de numerosos servicios en relación con los nueve factores importantes abordados anteriormente.

AVPreserve es una empresa de consultoría con servicio completo de archivos de medios y administración de datos. Nos asociamos con archivos, museos, organismos gubernamentales, corporaciones, medios y entretenimientos, y otras organizaciones que crean o recolectan medios para ayudarles a administrar, acceder y conservar sus valiosos activos y datos. Nuestros servicios abordan el ciclo de vida completo de las colecciones, desde la evaluación y planificación de preservación para materiales analógicos, incluyendo la administración de proyectos de actividades de digitalización, hasta los distintos aspectos de la preservación digital y administración de archivos, incluyendo la selección del monitoreo de actividad en bases de datos (DAM), el desarrollo de la taxonomía, las políticas y los flujos de trabajo, y el desarrollo de soluciones de software que apoyan la preservación y el acceso.

Esta obra está autorizada bajo licencia de Creative Commons Attribution-NonCommercial-NoDerivatives International 4.0 Para ver una copia de esta licencia, visite [http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en\\_US](http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en_US).