FACULTAD DE AGRONOMÍA – UNIVERSIDAD DE BUENOS AIRES

Políticas de Backups

Área de Comunicaciones, Seguridad y Nuevas Tecnologías – UTI - FAUBA

El presente documento tiene por objetivo definir las políticas de resguardo implementada en la Unidad de Tecnologías de la Información de la Facultad de Agronomía (UTI-FAUBA) de la Universidad de Buenos Aires, así como los servidores resguardados, los procedimientos de resguardo y recuperación, las personas implicadas y las posibles pérdidas bajo el presente plan.

Control de cambios del documento

Revisión	Fecha	Motivo	Autor
0	11/10/2011	Versión original del documento	Gustavo A. Marcello

Introducción

El presente documento especifica las políticas de backup para la UTI-FAUBA. Describe los servidores cuyos datos se resguardan, el personal autorizado a operar los backups y restauraciones, los datos a resguardar, el momento en que se ejecutan lo backups, el manejo de los datos almacenados, y las responsabilidades de los operadores y demás personas involucradas.

Alcance

Aplica a todos los datos alojados en servidores bajo resguardo acorde a los objetivos de seguridad de los datos establecidos por el área, y a los datos de usuarios que se incluyan (por mutuo acuerdo) en algún plan de backups brindado por el área.

Responsabilidades

Es responsabilidad de cada usuario conservar una copia de seguridad de todos sus archivos: Los usuarios son responsables por sus datos personales, como fotos, videos y archivos de música, ya sea que los mismos se encuentren en sus máquinas personales o en servidores de la UTI-FAUBA[1]. De la misma manera, es responsabilidad del usuario mantener siempre en su computadora o en otro medio alternativo, una copia (backup) de todos sus correos, archivos adjuntos, google docs, sites, calendars, etc. para poder reponerlos en caso de algún eventual problema en el servidor. La UTI-FAUBA no realiza copias de seguridad de las cuentas de Google-FAUBA, ya que las mismas son de carácter privado, y por otro lado no se cuenta con la infraestructura de almacenamiento suficiente para ello: actualmente existen 8000 cuentas Google-FAUBA, con una capacidad de 25,6 GB para cada una (solamente para los correos electrónicos), lo que resulta en una necesidad de 204800 GB (200 TB) de almacenamiento para realizar una sola copia de respaldo (un solo día) de todas las cuentas Google-FAUBA.

[1] Es un desperdicio de recursos de la institución resguardar datos que no son propios de ella y que no se necesitan para el normal funcionamiento de la misma. Se tendrá especial consideración en aquellos casos de archivos multimedia que formen parte del patrimonio institucional de la FAUBA.

Archivos que deben tener copias de respaldo

1. Servidor completo (incluyendo sistema operativo instalado) en el caso de servidores virtuales.

2. Software Base (Paquetes y/o Lenguajes de Programación con los cuales han sido desarrollados o interactúan nuestros Aplicativos Institucionales).

3. Software aplicativo (Considerando tanto los programas fuentes, como los programas objetos correspondientes, y cualquier otro software o procedimiento que también trabaje con los datos, para producir los resultados con los cuales trabaja el usuario final). Se deben considerar las copias de los listados fuentes de los programas definitivos, para casos de problemas.

4. Datos y estructura de datos (Bases de Datos, Índices, tablas de validación, contraseñas, tablespaces, usuarios, roles, configuraciones y todo archivo necesario para el funcionamiento de los Sistemas de Información de la Institución y la pronta recuperación de los mismos en caso de fallas).

5. Archivos de usuarios (Archivos utilizados por el personal docente y no docente de la FAUBA).

Especificaciones y normas para la elaboración de los backups en medios extraíbles

En caso de ser necesario, los backups deberán estar almacenados en DVD de primera marca, con etiquetas que contengan la siguiente información mínima:

- Nombre del Sistema y del Servidor.
- Versión / Año.
- Ubicación de los fuentes (en el caso de software aplicativo).
- Ubicación de los compilados (si corresponde, en el caso de software aplicativo).
- Ubicación de los instaladores (si corresponde, en el caso de software aplicativo).
- Número de DVD (en formato [número] de [cantidad total], ej. 2 de 3).

La ubicación de resguardo de dichos backups será la Sala de Servidores de Backup de la UTI, ubicado en el Pabellón de Botánica.

La frecuencia de obtención de éstos backups es anual en caso de que no se registren modificaciones; y cada vez que se elabore una nueva versión del sistema.

Criterios de elección de software para la gestión de backups

En este aspecto, en el mercado se dispone de un conjunto de utilitarios de propósito general que permiten efectuar procesos de respaldo de información, por lo tanto se debe disponer de criterios o metodologías apropiadas que permitan el uso de estos utilitarios de acuerdo a:

- Sistema operativo del origen.
- Sistema a resguardar y posibilidades de generar exportaciones de datos en tiempo real desde el mismo.
- Volumen de información a resguardar.
- Volumen de almacenamiento disponible en el servidor de backup y la posibilidad de comprimir los respaldos.
- Frecuencia de actualización de la información.
- Importancia de la información.

En los casos especiales en que no se disponga de un software apropiado para la generación del respaldo en cuestión, se deberá implementar un sistema confeccionado a medida de las necesidades particulares.

Metodología o procedimiento general de resguardo

Siempre que sea posible, el resguardo de los datos se hará de la siguiente forma:

- Todos los viernes a las 17:30 horas se hará un backup full de todos los servidores virtuales, incluyendo la instalación del sistema operativo.
- Todos los miércoles a las 0 horas se hará un backup full de todos los archivos críticos de los servidores.
- Todos los días (excepto los miércoles) se hará un backup diferencial de todos los archivos críticos de los servidores.

Bajo estas consideraciones se podrá recuperar una versión en particular para la última semana, además de las versiones completas. Ante la posible pérdida de archivos se podrá acudir a los backups diferenciales, y ante una pérdida total del servidor se podrá acudir a los respaldos full. De esta forma es posible recomponer el servidor completo, incluyendo la instalación del sistema operativo.

Procedimiento de restauración

La restauración se realiza siempre a un servidor de restauración administrado por el personal de la UTI-FAUBA. Ante la recuperación de datos personales, el responsable de los mismos deberá proporcionar una ubicación de red con los permisos adecuados de acceso y escritura o bien dirigirse a las oficinas de la UTI-FAUBA con un pendrive, DVD o disco portátil, en donde se le copiarán los datos que se hayan podido recuperar.

Ante la pérdida de datos, se deberá realizar una petición de recuperación lo antes posible, teniendo en cuenta que si transcurre mucho tiempo, el reúso de los volúmenes podría sobrescribir los datos resguardados.

Proceso de restauración mediante software Bácula

1. Para restaurar archivos desde la interfaz web de Bácula, dirigirse a Web - Restore
2. Seleccionar el cliente.
3. Elegir el respaldo que se desea restaurar.
4. Seleccionar el archivo que se desea restaurar, del árbol de archivos.
5. Elegir la versión del archivo (normalmente aparecerá sólo una) y lo arrastrar hacia la parte inferior de la pantalla.
6. Hacer clic en “Run”.
7. De ser necesario, seleccionar las opciones correspondientes y nuevamente hacer clic en “Run”.
8. Dirigirse a “main” y comprobar que el “job” está corriendo (o bien terminado).

Proceso de restauración mediante software SyncBack

1. Al iniciar, seleccionar el perfil del que se desea restaurar los archivos.
2. Hacer clic en el botón "Restaurar. Se recibirá un mensaje de advertencia, sobre la reescritura de los archivos en el origen. Hacer clic en “Yes”.
3. Si el mismo archivo existe en el origen y destino, entonces el comportamiento habitual del proceso es siempre sobrescribir el archivo de destino. Se puede cambiar este comportamiento a través de la pestaña "Avanzado" en las propiedades del perfil, para especificar una acción diferente (por ejemplo, para mantener siempre el archivo que se han modificado más recientemente).
4. Aparecerá una lista de archivos para los que la versión de la copia de seguridad difiere de la versión que actualmente existe en la ubicación que está la restauración de los archivos. Si las dos versiones del archivo son las mismas (existen y tienen el mismo nombre, fecha de modificación, tamaño, etc), entonces no están en la lista, porque no hay necesidad de restaurarlos.
5. Los dos casilleros de tipo “checkbox” que aparecen antes de cada nombre de archivo indican si el archivo en particular está presente en el origen (es decir, la copia de seguridad) y destino (es decir, la carpeta que se restaurará) respectivamente. El detalle de la columna “diferencia" explica la diferencia entre las dos versiones del archivo. Las columnas siguientes dan información sobre la fuente (copia de seguridad) y las versiones de destino del archivo.
6. Si se está restaurando una copia de seguridad completa (por ejemplo, por un fallo de disco), entonces no es necesario preocuparse demasiado acerca de este paso. Si se desea realizar la restauración de algunos archivos, es necesario asegurarse de que serán restauradas las versiones correctas de cada archivo.

7. Si se desea excluir determinados archivos de la restauración, hacer clic sobre ellos, y seleccionar "Skip file (do not copy)".
8. Cuando se esté satisfecho con la lista de archivos, hacer clic en "Continue Run" y los archivos serán restaurados desde la copia de seguridad.

Proceso de restauración de imágenes VMWare ESXi mediante ghettoVCB

1. Para este script no hay que realizar ninguna modificación en su interior, pero si se debe adecuar el archivo que indica las rutas desde donde ubicará la imagen y hacia donde se va a restaurar. El formato del archivo vm_a_restaurar debe ser: "Ruta_origen/imagen;/Ruta_destino;1". De manera detallada: toda la línea esta entre comillas separando los parámetros por ";". En el primer término se indica la ubicación de la imagen a restaurar. El segundo termino indica la ubicación donde se almacenan las VM's del ESXi. Y el último indica el formato de compresión de la imagen.
2. Ejecutar el script, con la siguiente orden: ~#ghettoVCB-restore.sh -c vm_a_restaurar -l /tmp/ghettoVCB-restore.log

Rutinas de seguimiento y control

Para prevenir fallas en la restauración de datos de servidores, la información almacenada en los backups debe ser verificada mensualmente y en forma íntegra. Para ello se deben realizar rutinas de control de backups.

Las mismas consisten en la prueba integral de los respaldos, desde su restauración desde el servidor de backup a otro de testeo (duplicado del servidor original en producción), descompresión e importación de datos hasta la pruebas correspondientes de funcionamiento (según el sistema que se esté recuperando). Ya que, por cuestiones de volúmenes de datos y tiempos, sería imposible realizar la comprobación de todos los backups de la FAUBA, se tomarán muestras de entre 5 y 10 backups aleatoria y rotativamente, para llevar adelante este proceso.

Posibles pérdidas de datos

Para que haya pérdida total de datos, deberán ocurrir los siguientes tres eventos simultáneamente:

- Pérdida de los datos en el servidor original, ya sea por negligencia o falla de hardware.
- Pérdida de los resguardos en el archivo en el/los servidor/es de backups, si se realizan a archivo[2].
- Pérdida de los datos en sistemas magnéticos o extraíbles, si se realiza este tipo de backup[3].

Recomendaciones

- Que todos los servidores tengan los datos importantes en particiones con RAID de nivel mayor que 0.
- Que se disponga de una red separada para la ejecución de los backups. De no ser posible, usar una VLAN para los backups.
- Que los servidores que resguarden un volumen grande de datos tengan interfaces de Gigabit.
- Que la red de backup esté a Gigabit. De no ser posible, por lo menos asegurar que ninguna de las interfaces por donde pasan los datos esté a menos de 100 mbps.
- Que los servidores destinados a backups, tengan fuente redundante.

Datos útiles de contacto

Nombre	Área	Teléfono	Mail	Oficina
Gustavo Marcello	CSyNT[4]	-	marcello@agro.uba.ar	UTI
Bruno Lottero	CSyNT	-	brunol@agro.uba.ar	UTI
David Vinazza	CSyNT	-	dvinazza@agro.uba.ar	UTI

Descripción de los servidores, servicios, datos y software utilizado

Ver tabla anexa

Terminología utilizada para la descripción:

Responsable del backup: Persona encargada realizar (en el caso de backups manuales) o programar y luego controlar las copias de seguridad. El responsable deberá estar claramente identificado y localizable en caso de incidencia urgente fuera del horario laboral

[2] Ya sea por rotura, negligencia o porque la política de backup haya especificado una rotación y reuso de los volúmenes. Además, siempre asumiendo que los datos se hayan resguardado con éxito en el medio de backup.

[3] Ídem punto anterior.

[4] Comunicaciones, Seguridad y Nuevas Tecnologías – UTI - FAUBA

Clasificación de la información. Por ejemplo en función de su importancia (crítica, importante o baja), en función de la sensibilidad de los datos de carácter personal tratados (alta, media, baja), etc. Esto nos permitirá considerar si las copias deben estar comprimidas, protegidas con contraseñas, o incluso cifradas.

Naturaleza de la información: Análisis de lo que se va a copiar de nuestros sistemas. ¿Qué tipo de información se va a copiar? Completo (clonación o imagen), Sistema (registros, configuración), aplicaciones, bases de datos, documentos, etc. En función de la naturaleza de la información será necesario tomar decisiones como por ejemplo el tipo de software de backup a utilizar para que permita "copias en caliente".

Volumen de información: Disponer de una estimación de la cantidad de datos a copiar. Esto será necesario para tener en cuenta por ejemplo que estrategia de copias utilizar, el tipo de soportes para realizar las copias o incluso estimar el tiempo necesario para realizar la copia.

Programación: La copia de respaldo se puede hacer de forma manual o automatizada y se debe tener en cuenta especialmente la hora elegida para hacerlas, prefiriendo las de menor actividad para reducir molestias a los usuarios (ya que no deberían estar trabajando mientras se hacen las copias). Ej. Horario nocturno.

Periodicidad/Frecuencia: La periodicidad calendárica es más sencilla de comprender y gestionar (diaria, semanal, mensual, anual). La frecuencia de las copias dependerá de los cambios producidos en la información y serán al menos semanales salvo que no se hayan producido modificaciones.

Tipo de backup: En función de la cantidad de información a copiar el backup puede ser Completo (toda), incremental (sólo se copian los ficheros modificados o creados desde la última copia incremental) o diferencial (copia los ficheros modificados desde la última copia completa). La incremental necesita menos espacio pero es más complicada de restaurar que la diferencial.

Localización/Almacenamiento: En función de la situación de las copias estas podrían ser locales si los soportes se almacenan en las mismas instalaciones en las que se encuentra el sistema de información, remotas si se almacenan en un pabellón diferente, y externas si son realizadas por internet y almacenadas en servidores externos a la facultad. Los almacenamientos locales deberían contar con la máxima seguridad física por ejemplo mediante el uso de armarios ignífugos bajo llave y en las correctas condiciones ambientales.

Soporte: Es el objeto físico que almacena o contiene datos o documentos, u objeto susceptible de ser tratado en un sistema de información y sobre el cual se pueden grabar y recuperar datos. Se debe elegir el tipo de soporte de almacenamiento que mejor se adapte a las necesidades particulares del sistema en cuestión.

Estrategia: Puede ser permanente o mediante la rotación de los soportes. Mientras más sencillo sea el esquema estratégico, más fácil de mantener. Por ejemplo hacer copias incrementales durante la semana de lunes a sábados y 1 completa los domingos, etc.

Software: La herramienta utilizada para realizar las copias debe soportar las características que se hayan planificado en función de las necesidades. Las copias se pueden programar con tareas mediante scripts de línea de comandos, o con un software específico (gratuito o invertir en sistemas de copias comerciales).

Pre-tareas: Acciones a realizar antes de realizar la copia de respaldo. Ej. Cerrar aplicaciones para una correcta copia, apagar servicios, etc.

Pos-tareas: Acciones a realizar al finalizar la copia de respaldo. Ej. Encender o volver a iniciar un servicio.

Origen: ¿Dónde se encuentra la información a copiar (maquina/ruta)? ¿Previsión de crecimiento?

Destino: ¿En qué soporte o máquina/ruta se harán las copias?

Control de copias/Supervisión: Es IMPORTANTÍSIMO verificar la correcta realización de las copias. Revisar logs, los soportes, y realizar pruebas de integridad de la información copiada es una parte fundamental del procedimiento.