Universitat Oberta de Catalunya

Implantación de Sistema de Backup Empresarial

Alberto Pozuelo Mozas
Administración de redes y sistemas operativos

Manuel Jesús Mendoza Flores

06/2017

Dedicatoria y agradecimientos

Este trabajo se lo dedico a mis padres, a mi mujer y a mi hermano que siempre me han estado apoyando durante toda mi vida. Sé que estas palabras no son suficientes para expresar mi agradecimiento, pero espero que con ellas, se den a entender mis sentimientos de aprecio y cariño a todos ellos.

FICHA DEL TRABAJO FINAL

Título del trabajo:	Implantación de Sistema de Backup
Nombre del autor:	Alberto Pozuelo Mozas
Nombre del consultor:	Manuel Jesús Mendoza Flores
Fecha de entrega (mm/aaaa):	06/2017
Área del Trabajo Final:	Administración de redes y sistemas operativos
Titulación:	*Grado de Ingeniería Informática - Tecnologías de la información*

Resumen del Trabajo (máximo 250 palabras):

Hoy en día, desde las empresas más pequeñas hasta las organizaciones más grandes, generan diariamente información de diversa índole cuyo valor es muy importante para la continuidad del negocio empresarial.

Dentro y fuera de dichas organizaciones, se producen eventos fortuitos o intencionados que comprometen dichos datos como, por ejemplo, desde inundaciones a sabotajes informáticos de los propios empleados.

La pérdida de dicha información puede comprometer no sólo, la continuidad del servicio, sino, también la continuidad del negocio, cuyos efectos serán muy negativos tanto para el empresario como para sus clientes.

En consecuencia, cualquier tipo de empresa necesita algún tipo de mecanismo que garantice la permanencia y recuperación consistente de los datos empresariales independientemente de la causa que lo ha provocado.

La implantación de un sistema de *backup* como el que trata este TFC garantizará al presente y futuro empresario no sólo la salvaguarda y recuperación de los datos, sino, la continuidad del servicio y del negocio.

Este Trabajo Final de Carrera analizará, desde el punto de vista de copias de seguridad, los problemas de la organización, las soluciones tecnológicas actuales y futuras (deduplicación y cloud), el proyecto de implantación de la solución y otros aspectos.

La forma de estructurar toda la información se basará en la metodóloga de proyecto PMBOK donde, por cada grupo de procesos se trabajarán a las áreas de conocimiento como, por ejemplo, dentro del proceso de Iniciación se desarrollará el Acta de Constitución de Proyecto.

Abstract (in English, 250 words or less):

Today, from the smallest companies to the largest organizations, daily generate information of various kinds whose value is very important for business continuity.

Within and outside such organizations, there are fortuitous or intentional events that compromise such data, for example, floods or computer sabotage by the employees themselves.

The loss of this information can not only compromise the continuity of the service, but also the continuity of the business, whose effects will be very negative for both the businessman and his clients.

Consequently any type of company needs some kind of mechanism that guarantees the permanence and consistent recovery of business data regardless of the cause that caused it.

The implementation of a backup system will guarantee the present and future businessman not only the safeguard and recovery of the data, but also the service continuity and business continuity.

This Final Career Work will analyze, from the point of view of backups, the problems of the organization, current and future technological solutions (deduplication and cloud), the project of implantation of the solution and other aspects.

The way to structure all the information will be based on the project methodology PMBOK where, for each group of processes will be worked to the areas of knowledge as, for example, within the process of Initiation, the Project Constitution Act will be developed.

Palabras clave (entre 4 y 8):

Garantía, permanencia, recuperación, consistencia del dato, continuidad del negocio, deduplicación y cloud.

Índice

Lista de tablas

Lista de ilustraciones

1. Introducción

1.1. Contexto y justificación del Trabajo

¿La empresa es consciente de todas sus vulnerabilidades?¿Qué es una empresa sin datos?¿Cómo debe actuar una empresa frente a la pérdida sus datos?¿Qué mecanismos pueden garantizar al empresario la recuperación consistente de sus datos (continuidad del negocio)?¿En cuánto tiempo puedo recuperar los sistemas y a partir de qué fecha y hora?

Cada día aparecen distintos tipos de amenazas que comprometen datos y archivos que las empresas generan para el desarrollo de su negocio. Estos tipos de amenaza varían desde errores humanos como, por ejemplo, el borrado accidental de cientos de ficheros hasta ataques informáticos como, por ejemplo, un malware que cifra los archivos y exige a la empresa (víctima) un recompensa económica a cambio del desbloqueo de los datos (ejemplo actual: *ransomware* criptográfico).

Es aquí donde las copias de seguridad toman un carácter relevante, destacado y de elevada importancia para la empresa pues, el empresario que dirige su negocio debe, de alguna forma, garantizar la continuidad de su negocio en caso de catástrofe. Además, no sólo se debe pensar en el empresario sino en todos los interesados (*steakholders*) que rodean la actividad empresarial donde cabe destacar el cliente.

Una buena infraestructura de *backup* responderá a garantías de integridad y consistencia del dato, a continuidad de servicio y de negocio. Con respecto a la continuidad del servicio, cabe destacar, dos objetivos fundamentarles de recuperación: tiempo máximo en el que se debe alcanzar un nivel de servicio mínimo tras una pérdida del servicio (RTO) y fecha y hora desde la que se restaurarán los datos tras una pérdida del servicio (RPO).

Como nota curiosa para aquellos que no lo saben, desde 2010, el 31 de Marzo se proclamó día mundial del *backup* a nivel cibernético. Véase los siguiente links.
http://cuantic.es/31-de-marzo-dia-mundial-del-backup/
http://www.worldbackupday.com/es/

1.2. Objetivos del Trabajo

Los objetivos del Trabajo son los siguientes:

- Tener una visión global de la necesidad de tener un sistema de *backup* en un entorno empresarial.
- Conocer las implicaciones que tiene la implantación de una arquitectura de *backup.*
- Conocer las últimas innovaciones a nivel *hardware* y *software* relacionados con el *backup.*
- Conocer las distintas estrategias de *backup* que puede implementar una empresa.
- Saber distinguir entre dato de negocio, dato de *backup* y archivado de la información.
- Saber definir SLA, RTO y RPO en base a la estrategia de *backup* elegida.
- Conocer las distintas estrategias de *backup* que puede implementar una empresa.

1.3. Enfoque y método seguido

Dependiendo de la naturaleza de la empresa, se puede enfocar estrategia para llevar a cabo la implantación de un entorno de *backup* de tres formas distintas:

a) Estrategia donde el empresario responsabiliza una **persona** (o a un conjunto de personas) la implantación y posterior gestión del *backup*.
b) Estrategia donde el empresario responsabiliza a una **empresa externa** la implantación y posterior gestión del *backup*.
c) Estrategia donde el empresario responsabiliza una **persona** (o a un conjunto de personas) **y** a una **empresa externa** la implantación y posterior gestión del *backup*.

En este caso, la estrategia más apropiada para conseguir los objetivos será aquella donde el empresario responsabiliza una persona (o a un conjunto de personas) y a una empresa externa la implantación y posterior gestión del *backup*.

Las ventajas de este tipo de estrategia son:

- Tener una persona in-situ que controle de forma diaria gestione los *backups*
- Tener un respaldo externalizado para que suplemente al administrador de *backup* en caso de baja o vacaciones.
- Tener un entorno actualizado y competente
- Garantizar SLA, RTO y RPO.
- Tener un centro de soporte donde resuelvan fallos de producto.
- Tener un proveedor que sea capaz de adaptarse y cumplir con los cambios de la compañía

Las desventajas de este tipo de estrategia son:

- Posible exceso de costes
- Diferencia de opiniones estratégicas de *backup*.

1.4. Planificación del Trabajo

La información que en este apartado se dividirá en dos partes:

1) Desde el punto de vista de implantación del entorno de backup

   A continuación se detalla la descripción de los recursos necesarios para realizar el trabajo y, posteriormente, un diagrama de Gantt donde se desglosa todas las fases, tareas, hitos y *steakholders* implicados para la implantación.

   Descripción de los recursos

   Los recursos se pueden dividir en dos grandes categorías:

   a. Recursos internos de la empresa

      Los recursos internos de la empresa los compone el administrador que gestione la herramienta de *backup* con responsabilidades de administrador, mientras que, el operador gestionará la herramienta de *backup* desde un punto de vista más operacional, es decir, monitorización de *backup*, cambio de cintas, ejecución periódica de *restores*, etc.

   b. Recursos externos de la empresa

      Los recursos externos de la empresa los compondrá el proveedor externo que realice la instalación, configuración y puesta en marcha del producto. En este caso, se han divido en dos tipos de recursos externos: proveedor comercial, aquel que gestionará la oferta desde un punto de vista económico y logístico y, proveedor técnico, aquello que trabajarán en la instalación y configuración del producto de *backup*.

Según el diagrama de Gantt será el administrador quien cubra el papel de jefe de proyecto porque estará siempre implicado en todas las tareas del proyecto y deberá reportar el avance del mismo. Este será el diagrama de Gantt del proyecto de implantación de *backup*.

Tabla 1.4-1 Diagrama de Gantt

Nombre	Duracion	Inicio	Fin	predecesores	Nombres del recurso
Implantación de solución de backup	40,778 days?	28/02/2017 8:00	25/04/2017 15:13		
Arranque	11,2 days	28/02/2017 8:00	15/03/2017 9:36		Administrador ;Proveedor-Comercial
Definición de estrategia, objetivos y alcance	1,6 days	28/02/2017 8:00	01/03/2017 13:48		Administrador [80%];Proveedor-Comercial[20%]
Análisis de la Solución	3,2 days	01/03/2017 13:48	06/03/2017 15:24	3	Administrador [80%];Proveedor-Comercial[20%]
Análisis de impacto y de riesgos	1,6 days	06/03/2017 15:24	08/03/2017 11:12	4	Administrador [80%];Proveedor-Comercial[20%]
Selección del software de backup	3,2 days	08/03/2017 11:12	13/03/2017 13:48	5	Administrador [80%];Proveedor-Comercial[20%]
Selección del implantador	0,8 days	13/03/2017 13:48	14/03/2017 11:12	6	Administrador [80%];Proveedor-Comercial[20%]
Acta de constitución del proyecto	0,8 days	14/03/2017 11:12	15/03/2017 9:36	7	Administrador [80%];Proveedor-Comercial[20%]
Hito: Aprobación y publicación del proyecto	0 days	15/03/2017 9:36	15/03/2017 9:36	8	
Planificación	21 days	15/03/2017 9:36	13/04/2017 9:36	9	
Establecer equipos de trabajo	2 days	15/03/2017 9:36	17/03/2017 9:36		Administrador [80%]
Definir y asignar roles y responsabilidades	2 days	15/03/2017 9:36	17/03/2017 9:36		Administrador [80%]
Cronograma y línea base	1 day	17/03/2017 9:36	20/03/2017 9:36	11;12	Administrador [80%]
Presupuesto	2 days	20/03/2017 9:36	22/03/2017 9:36	13	Administrador [80%]
Plan de gestión de riesgos	2 days	22/03/2017 9:36	24/03/2017 9:36	14	Administrador [80%]
Plan de calidad	2 days	24/03/2017 9:36	28/03/2017 9:36	15	Administrador [80%]
Plan de comunicaciones	2 days	28/03/2017 9:36	30/03/2017 9:36	16	Administrador [80%]
Plan de RRHH, administración y compras	2 days	30/03/2017 9:36	03/04/2017 9:36	17	Administrador [80%]
Establecer mapa de procesos y de datos	4 days	03/04/2017 9:36	07/04/2017 9:36	18	Administrador [80%]
Impacto y necesidades de infraestructura	4 days	07/04/2017 9:36	13/04/2017 9:36	19	Administrador [80%]
Hito: Aprobación y publicación de la planificación del proyecto	0 days	13/04/2017 9:36	13/04/2017 9:36	20	
Ejecución y Control	7,778 days	13/04/2017 9:36	24/04/2017 16:49	21	
Definir objetivos y alcance del sistema	0,889 days	13/04/2017 9:36	14/04/2017 8:42		Administrador [80%];Proveedor-Integrador
Definir requisitos en detalle	0,889 days	14/04/2017 8:42	14/04/2017 16:49	23	Administrador [80%];Proveedor-Integrador
Analizar intergración modulos del sistema y otras aplicaciones	1,333 days	14/04/2017 16:49	18/04/2017 10:29	24	Administrador [80%];Proveedor-Integrador
Implantación Backup	4,667 days	18/04/2017 10:29	24/04/2017 16:49	25	
Requisitos físicos de la solucion (alimentacion electrica, espacio en rack, etc.)	0,333 days	18/04/2017 10:29	18/04/2017 14:09		Administrador [50%];Proveedor-Integrador
Análisis, diseño y construcción de infraestructura de backup	1 day	18/04/2017 14:09	19/04/2017 14:09	27	Administrador [50%];Proveedor-Integrador
Instalación y parametrización del software de backup	2 days	19/04/2017 14:09	21/04/2017 14:09	28	Administrador [50%];Proveedor-Integrador
Pruebas y análisis de software de backup (backup y restore)	1 day	21/04/2017 14:09	24/04/2017 14:09	29	Administrador [50%];Proveedor-Integrador
Prueba de alta disponibilidad del software de backup	0,333 days	24/04/2017 14:09	24/04/2017 16:49	30	Administrador [50%];Proveedor-Integrador
Hito: Pruebas satisfactorias	0 days	24/04/2017 16:49	24/04/2017 16:49	31	
Cierre	0,8 days	24/04/2017 16:49	25/04/2017 15:13	32	
Entrega/recepción del Producto	0,2 days	24/04/2017 16:49	25/04/2017 9:25		Administrador ;Operador[50%];Proveedor-Integrador
Análisis de las lecciones aprendidas	0,2 days	25/04/2017 9:25	25/04/2017 11:01	34	Administrador ;Operador[50%];Proveedor-Integrador
Aprobación entregables	0,2 days	25/04/2017 11:01	25/04/2017 13:37	35	Administrador ;Operador[50%];Proveedor-Integrador
Acta de fin de proyecto	0,2 days	25/04/2017 13:37	25/04/2017 15:13	36	Administrador ;Operador[50%];Proveedor-Integrador
Hito: Cierre del Proyecto	0 days	25/04/2017 15:13	25/04/2017 15:13	37	
Gestión del Cambio	40,778 days	28/02/2017 8:00	25/04/2017 15:13		
Definir estrategia de gestión del cambio	2 days	28/02/2017 8:00	01/03/2017 17:00		
Comunicación	29,578 days	15/03/2017 9:36	25/04/2017 15:13		
Definir Plan de Comunicación	1,2 days	15/03/2017 9:36	16/03/2017 11:12	9	Administrador ;Operador[50%];Proveedor-Integrador
Presentación del proyecto	0,4 days	16/03/2017 11:12	16/03/2017 15:24	42	Administrador ;Operador[50%];Proveedor-Integrador
Comunicar avance del proyecto	8 days	16/03/2017 15:24	28/03/2017 15:24	43	Administrador ;Operador[50%];Proveedor-Integrador
Comunicar finalización del proyecto	0,8 days	24/04/2017 16:49	25/04/2017 15:13	32	Administrador ;Operador[50%];Proveedor-Integrador
Gestión de implicados (steakholders)	1,5 days	17/03/2017 9:36	20/03/2017 14:36	12	
Definir implicados del proyecto	1,5 days	17/03/2017 9:36	20/03/2017 14:36		Administrador [50%];Operador[50%];Proveedor-Comercial[50%];Proveedor-Integrador[50%]
Reunión inicial con implicados del proyecto	0,5 days	17/03/2017 9:36	17/03/2017 14:36		Administrador [50%];Operador[50%];Proveedor-Comercial[50%];Proveedor-Integrador[50%]
Gestión intereses, miedos, etc.	0,5 days	17/03/2017 9:36	17/03/2017 14:36		Administrador [50%];Operador[50%];Proveedor-Comercial[50%];Proveedor-Integrador[50%]
Documentación	16 days	15/03/2017 9:36	06/04/2017 9:36		
Definición de estándares de documentación	16 days	15/03/2017 9:36	06/04/2017 9:36	9	Proveedor-Comercial;Proveedor-Integrador;Administrador [50%]
Verificación documentación generada	16 days	15/03/2017 9:36	06/04/2017 9:36	9	Proveedor-Comercial;Proveedor-Integrador;Administrador [50%]
Formación	6,667 days	13/04/2017 9:36	21/04/2017 15:56	21	
Formación equipo de proyecto	0,667 days	13/04/2017 9:36	13/04/2017 15:56		Administrador [50%];Operador[50%];Proveedor-Integrador[50%]
Desarrollo de los contenidos formativos	5 days	13/04/2017 15:56	20/04/2017 15:56	54	Administrador [50%];Operador[50%];Proveedor-Integrador[50%]
Formación a usuarios	1 day	20/04/2017 15:56	21/04/2017 15:56	55	Administrador [50%];Operador[50%];Proveedor-Integrador[50%]
Gestión Proyecto	40 days?	28/02/2017 8:00	24/04/2017 17:00		
Reuniones de proyecto	40 days?	28/02/2017 8:00	24/04/2017 17:00		Administrador [50%];Proveedor-Integrador[50%];Operador[50%]
Monitorización y control	40 days?	28/02/2017 8:00	24/04/2017 17:00		Administrador [50%];Proveedor-Integrador[50%];Operador[50%]
Gestión de alcance y costes	40 days?	28/02/2017 8:00	24/04/2017 17:00		Administrador [50%];Proveedor-Integrador[50%];Operador[50%]
Gestión de incidencias y cambios	40 days?	28/02/2017 8:00	24/04/2017 17:00		Administrador [50%];Proveedor-Integrador[50%];Operador[50%]
Gestión de riesgos	40 days?	28/02/2017 8:00	24/04/2017 17:00		Administrador [50%];Proveedor-Integrador[50%];Operador[50%]
Gestión de la calidad del proyecto	40 days?	28/02/2017 8:00	24/04/2017 17:00		Administrador [50%];Proveedor-Integrador[50%];Operador[50%]
Información de Progreso	40 days?	28/02/2017 8:00	24/04/2017 17:00		Administrador [50%];Proveedor-Integrador[50%];Operador[50%]

2) Desde el punto de vista de TFG

   Se adjunta diagrama de Gantt donde se detallarán las tareas a realizar en este TFG.

Tabla 1.4-2 Diagrama de Gantt resumido

Actividad	Memoria	inicio	fin
Definir Plan de proyecto, y os implicados		20-feb	26-feb
Definir los objetivos y enfoque y método seguido en TFG	Indice	27-feb	05-mar
Entrega PEC1: Plan de Trabajo		06-mar	12-mar
Definición de estrategia, objetivos y alcance		13-mar	19-mar
Análisis de la Solución, plan necesidades de infraestructura		20-mar	26-mar
Análisis de impacto y de riesgos		27-mar	02-abr
Selección del software de backup		03-abr	09-abr
Entrega PEC2:	Background+prouesta	10-abr	16-abr
Requisitos físicos de la solucion (alimentacion electrica, espacio en rack, etc.)		17-abr	23-abr
Análisis, diseño y construcción de infraestructura de backup		24-abr	30-abr
Instalación y parametrización del software de backup		01-may	07-may
Pruebas y análisis de software de backup (backup y restore)		08-may	14-may
Entrega PEC3:	Resultados	15-may	21-may
		22-may	28-may
		29-may	04-jun
Entrega TFC y Presentación		05-jun	11-jun
		12-jun	18-jun
		19-jun	25-jun
Calificación final		26-jun	02-jul

1.5. Breve sumario de productos obtenidos

El producto principal que se obtendrá será una herramienta de *backup* para un entorno empresarial. Antes, durante y tras su finalización se obtendrá productos derivados de dicha actividad como, por ejemplo:

- Saber exactamente diferencial los modelos y estrategias de *backups* que puede adoptar en una empresa.
- Saber distinguir entre almacenamiento de dato de negocio, almacenamiento de archivado/historificación y almacenamiento de *backup*.
- Saber el coste aproximando de la implantación de un entorno de *backup*.
- Saber los beneficios estratégicos, tácticos y operaciones que aporta un entorno de *backup*.
- Saber las implicaciones adyacentes al entorno de *backup*: normativa legal.

1.6. Breve descripción de los otros capítulos de la memoria

A continuación, se describirá brevemente cada uno de los capítulos de la memoria del presente documento:

- Definición de estrategia, objetivos y alcance

  En este capítulo se describirá de definición de la estrategia, los objetivo a alto nivel y el alcance del proyecto de implantación de *backup*.

- Análisis de la solución, plan necesidades de infraestructura.

  En este capítulo se describirá de cómo se analizarán las distintas soluciones que hay en el mercado y, en base a cada una de las soluciones, las necesidades (aprovisionamiento) que serán necesarias adoptar por parte de la empresa.

- Análisis de impacto y de riesgos.

  En este capítulo se describirá el impacto que tiene en la organización la implantación de un sistema de *backup* y los riesgos que conlleva, no sólo su instalación y puesta en marcha, sino su mantenimiento.

- Selección del *software* de *backup*.

  En este capítulo se discutirá cual es el *software* de *backup* más adecuado determinados tipos de organizaciones empresariales.

- Requisitos físicos de la solución (alimentación eléctrica, espacio en rack, etc.).

  En este capítulo se analizar los requisitos físicos de la solución, como por ejemplo, tipo de alimentación eléctrica, climatización, etc .

- Análisis, diseño y construcción de infraestructura de *backup*.

  En este capítulo se desarrollará los tres pilares de la implantación de *backup* como son el análisis, el diseño y la forma instalar el producto.

- Instalación y parametrización del *software* de *backup*.

  En este capítulo se discutirá los requisitos previos al despliegue del

- Pruebas y análisis de *software* de *backup* (*backup* y *restore*).

  En este capítulo se analizará qué tipo de pruebas se deberán ejecutar para validar el correcto funcionamiento del entorno de *backup*.

1.7. Introducción a los riesgos y posibles contramedidas

En este punto, se describirá superficialmente los posibles riesgos y contramedidas tanto del proyecto como de la tecnología a implantar.

Desde el punto de vista de proyecto, pueden o no surgir problemas, por ello antes de empezar a con el proyecto se debe analizar los posibles riesgos y las posibles contramedidas.

Desde el punto de vista de la gestión de proyecto, podrían ser muchos pero cabe destacar: la falta de comunicación entre empresa e implantador, los problemas presupuestarios, los cambios en el alcance del proyecto, los accidentes, los problemas climáticos y la lentitud en la toma de decisiones.

Desde el punto de vista de la tecnología a implantar, se podrían encontrar muchos pero cabe destacar: incumplimiento de los requisitos del cliente, falta de estabilidad del producto, dependencia con otros productos y complejidad de gestión del producto.

Como contramedidas a los anteriores riesgos, se pueden utilizar un conjunto de herramientas y técnicas para gestionar los riesgos. Entre estas técnicas y herramientas cabe destacar un Plan de Gestión de Riesgos, un conjunto de definiciones de la probabilidad e impacto de los riesgos que se detallará en una matriz, un análisis de supuestos y un análisis DAFO (debilidades, amenazas, fortalezas y oportunidades).

Como consecuencia de los análisis de los riegos, se deberá realizar un análisis cualitativo y cuantitativo de los riesgos para garantizar que los objetivo marcados al principio del proyecto se cumplen. En el caso de los riesgos negativos o amenazas se deberá aplicar una de las siguientes estrategias: evitar, transferir, mitigar o, incluso, aceptar; mientras que, en el caso de los riesgos positivos u oportunidades se deberá aplicar una de las siguientes estrategias: explotar, compartir, mejorar o aceptar.

2. Entorno de *backup*

2.1. Definición de estrategia, objetivos y alcance

La definición de la estrategia para la implantación de un sistema de *backup* es una tarea muy difícil porque dependiendo de la naturaleza y de sus objetivos, ésta puede variar. No obstante, se propone el modelo organizativo basándose en el enfoque de los siguientes aspectos clave:

- Garantizar la coordinación y entendimiento entre todos los grupos de trabajo involucrados en el proyecto.
- Orientado a la optimización del servicio.
- Flujo de trabajo y responsabilidades perfectamente definidas.
- Apoyo de expertos en todas las áreas del proyecto.
- Equipos de trabajos flexibles para poder adaptarse y redefinirse según la evolución de las necesidades del proyecto para aprovechar todas las sinergias y garantizar los plazos y objetivos marcados.
- Mecanismos de escalado bien definidos.
- Comunicación formalizada a todos los niveles.
- Proporcionar información gradual del seguimiento del proyecto.

Por todo ello, los objetivos a alcanzar en la implantación de un entorno de *backup* radican en:

- Consolidar de forma segura las cargas de los aplicativos.
- Evitar el uso de recursos de infraestructura para tareas internas que impacten en el nivel de servicio.
- Un dimensionado de la solución basado en un estudio exhaustivo de los datos y necesidades reales, asegurando un rendimiento adecuado y predecible.
- Una estrategia de protección de la información completa y segura que garantiza el cumplimiento de los RPOs y RTOs en cada aplicativo.
- Asumir crecimientos impredecibles tanto en capacidad como en rendimiento sin tener que adquirir nuevos equipos o migrar, lo que representa una escalabilidad sin precedentes y una optimización de la inversión a futuro.

Todo aquello que no esté dentro de los objetivos y la estrategia anteriormente definida, estará fuera de alcance del proyecto donde, el acta de constitución será el testigo de dicho acuerdo.

2.2. Análisis de la solución, plan necesidades de infraestructura

Como paso inicial, se realizará el estudio de los escenarios de uso, para lo que se realizará la toma de datos de la situación actual del entorno actual. Adicionalmente, se mantendrán reuniones con los departamentos implicados en esta plataforma o servicio, con el fin de obtener toda la información relativa al mismo.

Incluye la recogida de la información para elaborar el diseño de la solución sobre la nueva infraestructura suministrada, a partir de la solución existente.

Así mismo incluye la recogida de la información detallada sobre la infraestructura de almacenamiento existente sus características técnicas y uso actual, ya sea en forma de inventarios, informes de estado, capacidad, distribución o rendimiento, que puedan ser interesantes para la ejecución del proyecto.

El mercado actual ofrece tres tipos de soluciones:

- **Solución on-premise:**

  Este tipo de solución implica que el cliente debe tener un espacio físico adecuado (en propiedad o alquilado) donde se pueda enchufar eléctricamente la solución de *backup*. Además, deberá cumplir con los requisitos legales actuales como por ejemplo la LOPD o PRL y otros requisitos como, por ejemplo, redundancia eléctrica, protección antiincendios, control de acceso a sala, etc.

  La forma física de los dispositivos de *backup* se puede dividir en:

  - **Cabina física de *backup* con cintas**: el dispositivo final donde se guardarán los datos serán cintas físicas cuyos modelos puede variar desde DDS a LTO6

  - **Cabina física de *backup* con discos**: el dispositivo final donde se guardarán los datos serán cintas virtuales cuyos modelos puede variar desde DDS a LTO6 dentro de cabinas virtuales (VTL)

- **Solución *cloud***

  Este tipo de solución implica que el cliente debe tener un contrato con algún proveedor de servicios en la nube. Además, dependiendo del proveedor esté contratado, el producto podrá variar (competencia entre distintos proveedores de *cloud* y *backup*). Igualmente, deberá cumplir con los requisitos legales actuales como por ejemplo la LOPD o PRL y, sobre todo, la geolocalización de los datos salvaguardados.

  No existe forma física de los dispositivos de *backup* sino que tomaría las mismas características que una librería virtual con sus cintas virtuales pero gestionando el *backup* de forma virtual.

- **Solución hibrida/mixta**

  Este tipo de solución mezcla la tecnología de las dos anteriores, es decir, se puede hacer *backup* a cinta o disco físico y otros *backups* se salvaguardaran en la nube del proveedor. Si el *backup* se realiza a librería física, se obtendrá mayor rendimientos mientras que si se realiza a nube, el rendimiento será menor.

La clasificación anterior se debe cruzar con las necesidades de la empresa. Dichas necesidades pueden ser:

a) Garantizar la continuidad del negocio y del servicio (RPO y RTO).
b) Alinearse con la estrategia de negocio y los procesos de negocio.
c) Alinearse con la legislación del país.
d) Ahorrar costes con respecto al entorno actual.
e) Mejorar el rendimiento de *backup* y recuperación de datos.

En la actualidad, se está tendiendo a elegir una tecnología híbrida donde el cliente exige:

I. que determinados *backups* sean recuperados lo más rápido posible y en un punto en el tiempo, ara que no impacte económicamente en su negocio (y por ende a su cliente)
II. que el proceso de copia se ejecute lo más rápido posible para que no colisionen con los procesos de negocio.
III. que los datos estén salvaguardados durante varios años debido al nivel de seguridad del propio dato exigido por la normativa del país.
IV. ahorro de costes quitando dentro de su presupuesto económico la compra de cintas físicas (y los costes asociados al vaulting si los tuviera)

2.3. Análisis de impacto y riesgos

Por definición, un riesgo es un evento o condición incierta que, si sucede, tiene un efecto en por lo menos uno o varios de los objetivos del proyecto (alcance, cronograma, coste y calidad). En este caso, se pueden identificar determinados riesgos previsibles pero, a lo largo del proyecto, podrán aparecer otros riesgos (condición de incertidumbre).

Desde el punto de vista de la **gestión de proyecto**, se podrían identificar los siguientes riesgos y contramedidas:

- **Falta de comunicación entre empresa e implantador**: la comunicación es el hilo conductor y principal entre el cliente y el implantador. Si este hilo se corta, ambos no tendrán una perspectiva actualizada de la evolución del proyecto ni de la opinión (satisfactoria o no) del proyecto.

  Como contramedida a este riesgo, se deben establecer reuniones periódicas de seguimiento de proyecto y canales adicionales de comunicación donde poder expresar los problemas y las soluciones encontradas.

- **Lentitud en la toma de decisiones**: a lo largo de un proyecto pueden surgir cambios inesperados que afecten al ritmo de trabajo del proyecto donde la velocidad toma de la toma de decisiones, afectará a la planificación del proyecto (provocando un desvío).

  Como contramedida a este riesgo, en la fase de planificación se realizará un análisis cualitativo de los posibles riesgos donde una matriz de probabilidad e impacto categorizará los riesgos y se establecerán las prioridades de los riesgos.

- **Problemas presupuestarios**: los costes asociados al proyecto pueden verse afectados por diversos factores como, por ejemplo, un fallo en la estimación de duración de las actividades (a mayor duración de la actividad, mayor coste) o por la falta de definición de la actividad como, por ejemplo, si un técnico tuviera que hacer la actividad A y finalmente se decide realizar la actividad B, la definición de la actividad tendrá un coste adicional.

  Como contramedida a este riesgo, en la fase de planificación, se deben analizar y definir correctamente tanto el tipo de actividad como la duración de la propia actividad.

- **Cambios en el alcance del proyecto**: al principio de proyecto, en el documento de alcance de proyecto, se identificaron los supuestos del proyecto. Si se varían dichos supuestos, la línea base del proyecto se modificará cuya consecuencia será un desvío en el proyecto. El desvío en la línea base puede afectar en mayor o

menor medida al proyecto pero es una causa potencial de riesgo cuya consecuencia puede llegar a provocar el fracaso del proyecto.

Como contramedida a este riesgo, se debería tener firmado (tanto por el cliente como por el implantador) el documento de alcance de proyecto y recordar en las reuniones periódicas, los supuestos del proyecto.

- **Accidentes y problemas climáticos**: existen factores o agentes externos al proyecto que pueden ser causas potenciales de riesgo como, por ejemplo, inundaciones lluvias torrenciales o tsunami, terremotos de elevada intensidad o simplemente accidentes humanos que perjudican al proyecto.

  Como contramedida a este riesgo, se debería realizar un análisis de las listas de control donde la información histórica y el conocimiento acumulado prevean posibles accidentes.

Desde el punto de vista de la **tecnología a implantar** se podrían identificar los siguientes riesgos:

- **Incumplimiento de los requisitos del cliente**: entender los requisitos de cliente sin suponer un aspecto obvio es fundamental para saber si la tecnología de *backup* a implantar es la más adecuada. Además, es muy importante que el cliente conozca las limitaciones del producto de *backup* antes de adquirirlo, no dejando margen de duda en cualquiera de los aspectos que puedan impactar en la continuidad del servicio o del negocio. Por ejemplo, si el producto de *backup* no cubre la necesidad de hacer *backup* un motor de base de datos Informix de forma caliente pero sí puede hacer *backup* de dicha base de datos estando parada, tanto el cliente como el implantador deben reflejar claramente su comprensión porque, si no se puede parar la base de datos de Informix por continuidad del servicio, no se cumplirá con el requisito del cliente; mientras que si el cliente asume un *backup* frio de la base de datos, el implantador debe reflejar claramente las jornadas extras del desarrollo de un script que automatice tanto la parada como el arranque de dicha base de datos.

  Como contramedida a este riesgo, el implantador debe realizar varias sesiones informativas o presentaciones del producto donde, el cliente confirme tanto su comprensión tecnológica como su impacto a nivel de negocio.

- **Falta de estabilidad del producto**: al inicio de la instalación todo entorno de *backup* suele aparentar una estabilidad sólida pero, según pasa el tiempo, el almacenamiento y la gestión de datos salvaguardo crece de forma considerable pudiendo aparecer determinada inestabilidad. Dicha inestabilidad se podría manifestar de diferentes formas como, por ejemplo, pérdida de información del catálogo de objetos copiados, ralentización de las copias de seguridad por falta de indexación en el catálogo de objetos copiados, falta de liberación de espacio (*housekeeping*) debido a incoherencia entre dato vigente y dato expirado, etc. Todas estas razones hacen que un producto de *backup* pase a ser catalogado un entorno inestable donde la garantía de salvaguarda de datos y recuperación deje de existir poniendo en riesgo la continuidad del servicio y del negocio.

  Como contramedida a este riesgo, el cliente detallará qué, cuándo, cómo, dónde, porqué y de que se quiere hacer *backup*/*restore* mientras que, el implantador deberá definir tanto los mantenimiento del producto como las políticas de *backup* y *restore* en base a lo descrito por el cliente.

- **Compatibilidad**: las presentaciones que los comerciales y comerciales técnicos realizan en el cliente suelen ser sobre una plataforma ofimática donde todo lo que se explica funciona. La realidad suele ser otra porque el producto que se ha instalado convive con un conjunto de elementos y variables ajenas al *power point*. Por ejemplo, si el almacenamiento de la nueva herramienta de *backup* se conecta a través de cables de fibra (dentro de una SAN) es imprescindible saber que los niveles de firmware de dicha cabina son compatibles tanto con los switches de fibra como con el resto de servidores adyacentes. La incompatibilidad de hardware de *backup* con el resto de dispositivos tendrá como consecuencia o bien, la incapacidad de arrancar el almacenamiento de *backup*, o bien, someter al resto de dispositivos a una actualización del firmware (dicha actualización tendrá impacto en la empresa porque requiere corte de servicio).

  Como contramedida a este riesgo, el cliente tendrá que entregar documentación detallada de toda su infraestructura y el implantador tendrá que trazar una matriz de compatibilidad para asegurar que todos los dispositivos adyacentes al *backup* son compatibles.

- **Complejidad de gestión del producto**: la usabilidad de una herramienta de *backup* y su curva de aprendizaje puede suponer un riesgo considerable para la empresa. Si el *software* de *backup* es muy difícil de gestionar y para recuperar una base de datos se deben realizar muchas tareas, la duración de la recuperación incrementará impactando en el servicio e incluso en el negocio. Además, no se cumplirán con los acuerdos de nivel de servicio (ANS/SLA) ni con los tiempo de recuperación objetivo (RTO) donde se establece un tiempo objetivo para la reanudación de sus servicios después de un desastre.

Como contramedida a este riesgo, el cliente deberá evaluar los productos de *backup* en función de las personas que lo vayan a gestionar. Así mismo, el implantador deberá definir una guía básica (o de supervivencia) ante posibles desastres.

2.4. Selección del *software* de *backup*

Dentro de cada una de las aplicaciones de *backup*, se deben analizar las características esenciales y comunes a cada uno de ellos. Dichas características serán los pilares comparativos entre las distintas aplicaciones de *backup* que hay en el mercado. Su descripción es la siguiente:

a) Garantía de copia y recuperación consistente

   Todo *software* de *backup* debe garantizar que puede salvaguardar todos los datos críticos para el negocio y que dichos datos pueden recuperarse en, al menos, el mismo entorno donde estaban de forma consistente.

b) Garantía de continuidad de servicio y negocio

   Todo *software* de *backup* debe garantizar que, aunque haya un problema con el catálogo de datos de la herramienta de *backup*, es capaz de reconstruirse o recuperarse sin perder ninguna referencia de los *backup* ya realizados.

c) Garantía de seguridad de la información

   Todo *software* de *backup* debe garantizar que la confidencialidad, la integridad y la autenticación de los datos se mantienen en las copias de seguridad.

d) Garantía de compatibilidad con aplicaciones de negocio

   Todo *software* de *backup* debe garantizar que es compatible con las aplicaciones críticas para la continuidad del negocio. Por ejemplo, si una empresa tiene como base de datos Oracle, la herramienta de *backup* debe ser compatible con dicha base de datos para hacer *backup*.

e) Garantía de almacenamiento

   Todo *software* de *backup* debe garantizar que se puede salvaguardar los datos en al menos un tipo almacenamiento (cinta, disco o nube).

f) Garantía de automatización y retención de la información

   Todo *software* de *backup* debe garantizar se puede automatizar tanto las copias de seguridad como las recuperaciones y que los datos salvaguardado con una protección no son eliminados o sobrescritos.

g) Garantía de usabilidad e informes

Todo *software* de *backup* debe garantizar que la gestión de su herramienta es inuititva y ágil. Además debe ser capaz de ofrecer informes y estadísticas de *backup.*

Antes de empezar a detallar cada uno de las aplicaciones de *backup* que hay en el mercado, cabe destacar que debido al extenso mercado sólo se han elegido los cuatro primeros *software* de *backup* con mejor clasificación dentro del cuadrante mágico de Gartnet (*software* propietario) o con mejor reputación dentro de las comunidades Opensource (*software opensource*).

2.4.1. *Software* de *backup Opensource*

AMANDA (ZMANDA)

1) Que es Amanda (http://wiki.zmanda.com/index.php/Main_Page)

Amanda es una solución de copia de seguridad opensource que permite al administrador de *backup* configurar un servidor central (quien gestionará todas la copias) para realizar copias de seguridad de distintos servidores clientes o estaciones de trabajo (portátiles) a través de la red a unidades de cinta, discos o a la nube.

Amanda aprovecha herramientas o utilidades nativas de los propios sistemas operativos o aplicaciones (por ejemplo, dump o tar). Puede ejecutar copias de seguridad de un gran número de servidores y estaciones de trabajo que basada en diferentes versiones de Linux o Unix. Además, aprovecha el cliente nativo de Windows para realizar copias de seguridad de escritorios y servidores de Microsoft Windows.

2) Historia

Originalmente fue escrito por James da Silva mientras que en la Departamento de Ciencias de la Computación de la Universidad de Maryland fue donde se empezó a desarrollar y probar. Con el tiempo, Blair Zajac recogió la versión 2.3.0 y asumió la tarea de convertir a Amanda para usar el sistema GNU para la configuración. Después, se creó un equipo básico de desarrollo de Amanda cuya estabilidad no ha sido muy sólida, viendo a desarrolladores entrar sy salir del proyecto.

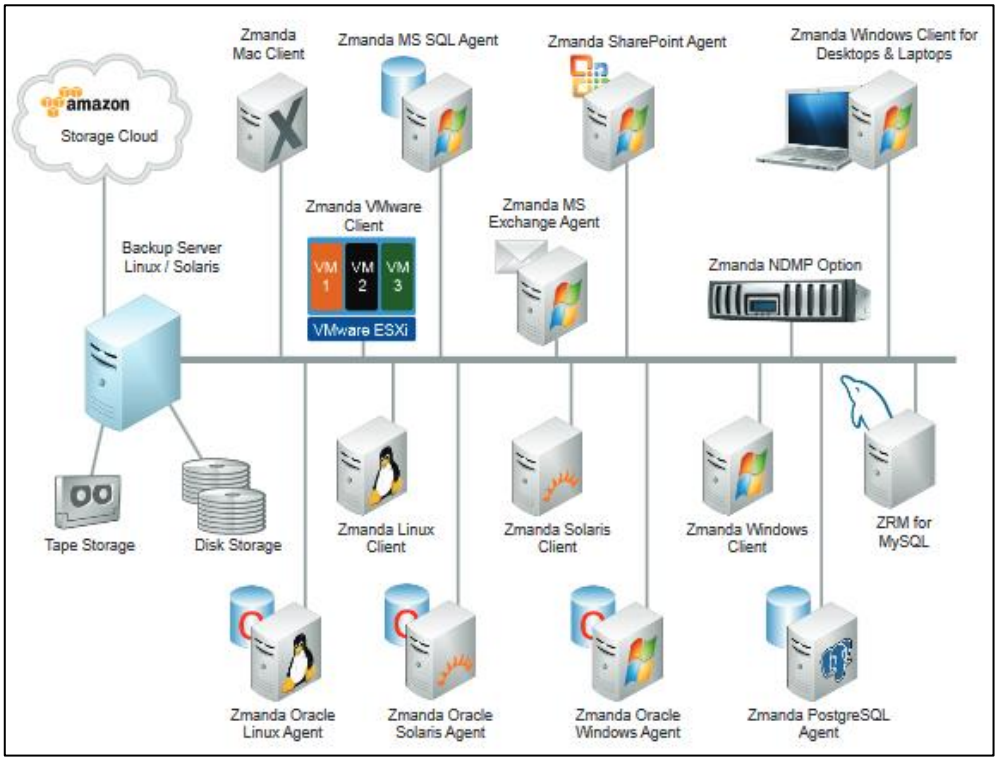
Después de que el mantenimiento dejó de ser apoyado por el Departamento de Ciencias de la Computación, AMANDA movió su repositorio CVS y su sitio web principal a Sourceforge, donde está alojado en la actualidad.

A día de hoy, el desarrollo de Amanda es apoyado por Zmanda.

3) Características esenciales y comunes

A continuación se presenta un esquema de la arquitectura de *backup* con distintos clientes

Ilustración 2.4.1-1 Arquitectura Zmanda



wiki.zmanda.com. (2017). *Overview - Introduction*. [online] Disponible en: http://www.zmanda.com/amanda-enterprise-edition.html [Accedido 21 de Marzo del 2017].

Antes de empezar a describir las características esenciales, cabe destacar que el *software* de *backup* Amanda se ha divido en dos grande bloques: aquel soportado por la Comunidad y aquel soportado a nivel empresarial. En este caso se analizará sólo el soportado por la Comunidad (se ajdunta cuadro de diferencias).

Tabla 2.4.1-1 Diferencias Community y Entrerprise

Feature	Community	Enterprise
Centralized Full and Incremental Backups	SI	SI
Linux and UNIX Support	SI	SI
Windows Server and Desktop Support	SI	SI

Mac OS X Support	SI	SI
Intelligent Backup Scheduler	SI	SI
Backup to Disk (NAS, SAN, iSCSI, Cloud Storage)	SI	SI
Backup to Tape Drives, Tape Libraries and VTLs	SI	SI
Vaulting & Disk-to-Disk-to-Tape (D2D2T)	SI	SI
Open Formats for long-term archiving	SI	SI
Encryption and Compression of backup archives	SI	SI
Certification with Security Enhanced Linux (SELinux)	NO	SI
Live Backup of Oracle	NO	SI
Live Backup of SQL Server, Exchange and SharePoint	NO	SI
Backup of images of live VMware based VMs	NO	SI
NDMP based backups of NAS Appliances	NO	SI
Web Based Management Console	NO	SI
Backup Reporting	NO	SI
Role-based Administration	NO	SI
Wizard-driven Installer	NO	SI
Backup Server Replication (DR to a remote site)	NO	SI
24x7 Production Support	NO	SI
Professional Services and Training	NO	SI

wiki.zmanda.com. (2017). *Comparison of Amanda Community and Enterprise Edition*. [online] Disponible en: http://www.zmanda.com/Amanda-Enterprise-Amanda-Community-comparison.html [Accedido 11 de Marzo del 2017].

Tabla 2.4.1-2 Características esenciales de Amanda

Producto/Fabricante	Garantía de copia y recuperación consistente	Garantía de continuidad de servicio y negocio	Garantía de seguridad de la información	Garantía de compatibilidad con aplicaciones de negocio	Garantía de almacenamiento	Garantía de automatización y retención de la información	Garantía de usabilidad e informes
Amanda	Amanda cumple con la garantía de copia y recuperación consistente de todos los datos críticos para el negocio y que dichos datos pueden recuperarse en, al menos, el mismo entorno donde estaban de forma consistente	Amanda cumple con la garantía de continuidad de servicio y negocio porque el es capaz de reconstruirse o recuperarse sin perder ninguna referencia de los *backup* ya realizados	Amanda cumple con la garantía de seguridad de la información porque es capaz de implentar algoritmos de cifrados asimétricos.	Amanda cumple parcialmente con la garantía de compatibilidad de aplicaciones de negocio. Faltan más compatibilidad con otras aplicaciones de mercado.	Amanda cumple con la garantía de almacenamiento porque es capaz de salvaguardar los datos en al menos un tipo almacenamiento (cinta, disco o nube).	Amanda cumple con la garantía de automatización y retención de la información porque se puede automatizar tanto las copias de seguridad como las recuperaciones y proteger los datos con retención	Amanda no cumple con la garantía de usabilidad e informes porque no dispone de una interfaz gráfica intuitiva y ágil. Tampoco tiene un apartado especial para informes

BACKUPPC

1) Que es BackupPC (http://backuppc.sourceforge.net/)

   BackupPC es una solución de copia de seguridad opensource que proporciona a la hora de realizar copias de seguridad alto rendimiento y de calidad empresarial. Es compatible con PCs, escritorios y portátiles Unix, Linux, Windows y MacOS, y almacena las copias en el disco de un servidor.

   BackupPC tiene herramientas o utilidades que hacen que minimice el almacenamiento en disco y la E/S de disco. Esto es así porque los archivos idénticos de diferentes copias de seguridad se almacenan sólo una vez (utilizando enlaces). No es necesario ningún cliente, ya que el propio servidor es un cliente para varios protocolos que son manejados por otros servicios nativos del sistema operativo cliente.

2) Historia

   Originalmente fue escrito por Craig H. Barratt sobre el 2001. A partir de entonces distintos desarrolladores han contribuido a mejorar esta aplicación proclamándose en 2007 una de las 3 herramientas de *backup* opensource más conocidas.

3) Características esenciales y comunes

Tabla 2.4.1-3 Características esenciales de BackupPc

Producto/Fabricante	Garantía de copia y recuperación consistente	Garantía de continuidad de servicio y negocio	Garantía de seguridad de la información	Garantía de compatibilidad con aplicaciones de negocio	Garantía de almacenamiento	Garantía de automatización y retención de la información	Garantía de usabilidad e informes
BackupPc	BackupPC cumple con la garantía de copia y recuperación consistente de todos los datos críticos para el negocio y que dichos datos pueden recuperarse en, al menos, el mismo entorno donde estaban de forma consistente.	BackupPC cumple con la garantía de continuidad de servicio y negocio porque el es capaz de reconstruirse o recuperarse sin perder ninguna referencia de los *backup* ya realizados.	BackupPC cumple con la garantía de seguridad de la información porque es capaz de implentar algoritmos de cifrados asimétricos.	BackupPC cumple no con la garantía de compatibilidad de aplicaciones de negocio porque sólo realiza *backup* de ficheros.	BackupPC cumple con la garantía de almacenamiento porque es capaz de salvaguardar los datos en al menos un tipo almacenamiento (cinta, disco o nube). En este caso no se integra con cloud.	BackupPC cumple con la garantía de automatización y retención de la información porque se puede automatizar tanto las copias de seguridad como las recuperaciones y proteger los datos con retención	BackupPC cumple con la garantía de usabilidad e informes porque dispone de una interfaz gráfica intuitiva y ágil. También tiene un apartado especial para informes.

BACULA

1) Que es Bacula (http://blog.bacula.org/)

Bacula es una solución de copia de seguridad opensource que permite que un administrador del sistema gestione la copia de seguridad, la recuperación y la verificación de los datos de la computadora a través de una red de computadoras de diferentes tipos. Bacula también puede ejecutarse completamente en una sola computadora (aunque puede que no sea la mejor opción para este tipo de servicio) y puede hacer copias de seguridad de varios tipos de medios, incluyendo cinta y disco.

Su infraestructura está basada en cliente/servidor. Bacula es relativamente fácil de usar y eficiente, a la vez que ofrece muchas funciones avanzadas de administración de almacenamiento que facilitan la búsqueda y recuperación de archivos perdidos o dañados. Es compatible con PCs, escritorios y portátiles Unix, Linux, Windows y MacOS, y almacena las copias en el disco de un servidor.

2) Historia

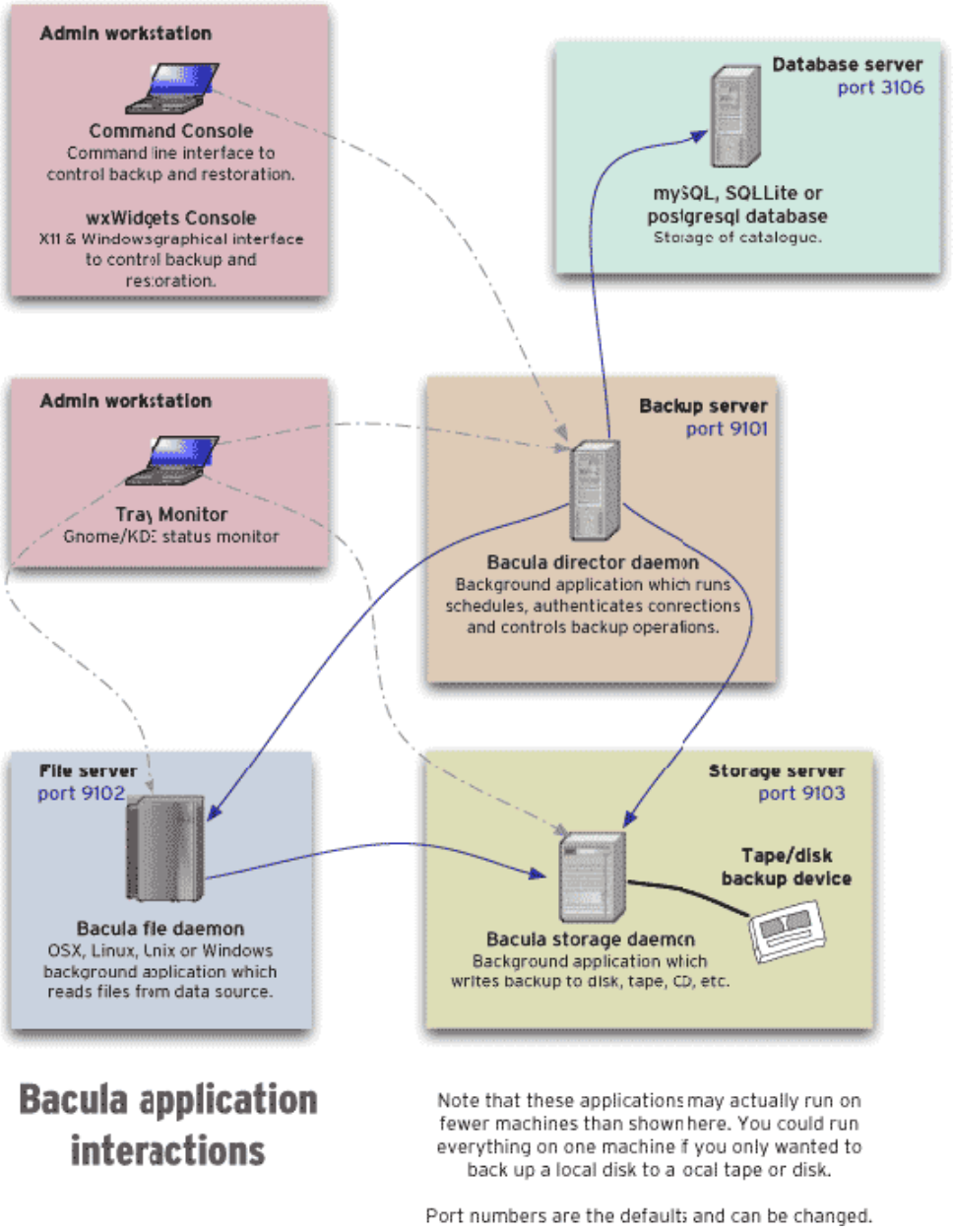
El proyecto de Bacula comenzó en el año 2000. En 2011, Graham Keeling, un "ex" desarrollador de la comunidad Bacula, propuso una nueva línea de desarrollo y en 2013 fue autorizado por Bacula Systems para usar parte de código en su nueva líneas. Ese mismo año, en el mes de febrero, un ex desarrollador de la comunidad Bacula (con varios otros usuarios de *software* libre) lanzó Bareos (se verá en el siguiente punto).

http://wiki.bacula.org/doku.php?id=bacula_manual:what_is_bacula

3) Características esenciales y comunes

A continuación se presenta un esquema de la arquitectura de *backup* con distintos clientes

Ilustración 2.4.1-2 Arquitectura Bacula



wiki.bacula.org. (2017). *Bacula Components or Service*. [online] Disponible en: http://wiki.bacula.org/lib/exe/detail.php?id=bacula_manual%3Awhat_is_bacula&media=bacula_manual:bacula-applications.png [Accedido 22 de Marzo del 2017].

Antes de empezar a describir las características esenciales, cabe destacar que el *software* de *backup* Bacula se ha divido en dos grande bloques: aquel soportado por la Comunidad y aquel soportado a nivel empresarial. En este

caso se analizará sólo el soportado por la Comunidad (se adjunta cuadro de diferencias).

Tabla 2.4.1-4 Diferencias entre Community y Enterprise

Advanced Features	Community	Enterprise
Snapshot technology	NO	SI
Snapshot management	NO	SI
Single file restore for VMware	NO	SI
Single mailbox recovery for Exchange	NO	SI
Client initiated backup	NO	SI
Windows SD	NO	SI
Windows EFS support	NO	SI
Storage device switchover	NO	SI
Restart failed job	SI	SI
Communication line statistics	NO	SI
Catalog performance improvement	NO	SI
Periodic statistics for running jobs in Director	SI	SI
Truncate command	SI	SI
SD to SD replication	SI	SI
SD to SD replication with Deduplication	NO	SI
Communication Line Compression	NO	SI
Readonly drive directive	SI	SI
Catalog schema for high performance	NO	SI
Global Endpoint Deduplication	NO	SI
Aligned Volume Format	NO	SI
Plugins	**Community**	**Enterprise**
LDAP and Active Directory	NO	SI
VMware	NO	SI
KVM	NO	SI
Hyper-V	NO	SI
Plugin for SAP	NO	SI
Windows VSS	NO	SI
Windows Bare Metal Recovery	NO	SI
Linux Bare Metal Recovery	NO	SI
NDMP Plugin	NO	SI
Incremental Accelerator for Netapp plugin	NO	SI
Oracle plugin with SBT	NO	SI
PostgreSQL plugin	NO	SI
MySQL plugin	NO	SI
Plugin for MSSQL Server	NO	SI
Delta plugin	NO	SI
San Shared Storage plugin	NO	SI

www.baculasystems.com. (2017). *Bacula Community version compared with Enterprise Edition 8*. [online] Disponible en: https://www.baculasystems.com/products/selective-migration-plan/enterprise-community-comparison [Accedido 22 de Marzo del 2017].

Tabla 2.4.1-5 Características esenciales de Bacula

Producto/Fabricante	Garantía de copia y recuperación consistente	Garantía de continuidad de servicio y negocio	Garantía de seguridad de la información	Garantía de compatibilidad con aplicaciones de negocio	Garantía de almacenamiento	Garantía de automatización y retención de la información	Garantía de usabilidad e informes
Bacula	Bacula cumple con la garantía de copia y recuperación consistente de todos los datos críticos para el negocio y que dichos datos pueden recuperarse en, al menos, el mismo entorno donde estaban de forma consistente.	Bacula cumple con la garantía de continuidad de servicio y negocio porque él es capaz de reconstruirse o recuperarse sin perder ninguna referencia de los *backup* ya realizados.	Bacula cumple con la garantía de seguridad de la información porque es capaz de implentar algoritmos de cifrados asimétricos.	Bacula cumple parcialmente con la garantía de compatibilidad de aplicaciones de negocio. Faltan más compatibilidad con otras aplicaciones de mercado.	Bacula cumple con la garantía de almacenamiento porque es capaz de salvaguardar los datos en al menos un tipo almacenamiento (cinta, disco o nube). En este caso no se integra con cloud.	Bacula cumple con la garantía de automatización y retención de la información porque se puede automatizar tanto las copias de seguridad como las recuperaciones y proteger los datos con retención.	Bacula cumple con la garantía de usabilidad e informes porque dispone de una interfaz gráfica intuitiva y ágil. También tiene un apartado especial para informes.

BAREOS

1) Que es Bareos (https://www.bareos.org/en/)

Bareos es una solución de copia de seguridad opensource muy parecida a Bacula. Es decir, un administrador del sistema gestionará la copia de seguridad, la recuperación y la verificación de los datos de la computadora a través de una red de computadoras de diferentes tipos. También puede ejecutarse completamente en una sola computadora (aunque puede que no sea la mejor opción para este tipo de servicio) y puede hacer copias de seguridad de varios tipos de medios, incluyendo cinta y disco.

Su infraestructura está basada en cliente/servidor. Bareos es relativamente fácil de usar y eficiente, a la vez que ofrece muchas funciones avanzadas de administración de almacenamiento que facilitan la búsqueda y recuperación de archivos perdidos o dañados. Es compatible con PCs, escritorios y portátiles Unix, Linux, Windows y MacOS, y almacena las copias en el disco de un servidor.

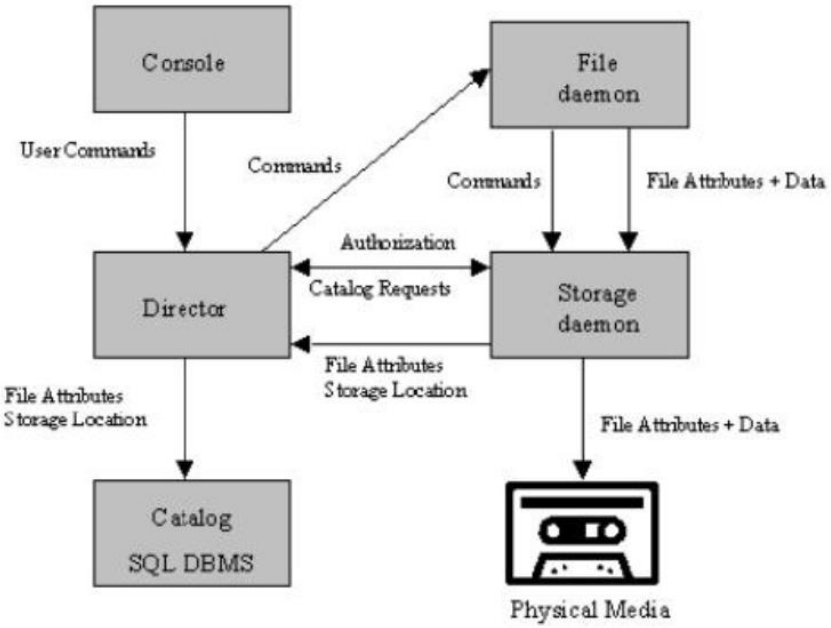
2) Historia

Desde el 2010, Bareos sigue su propio camino haciendo la competencia a Bacula desde dos tipos de versiones: versión opensource y versión empresarial.

3) Características esenciales y comunes

A continuación se presenta un esquema de la arquitectura de *backup* con distintos clientes

Ilustración 2.4.1-3 Arquitectura de Bareos



doc.bareos.org. (2017). *Interactions Between the Bareos Services*. [online] Disponible en: http://doc.bareos.org/master/html/bareos-manual-main-reference.html#x1-50001.2 [Accedido 27 de Marzo del 2017].

Antes de empezar a describir las características esenciales, cabe destacar que el *software* de *backup* Bareos se ha divido en dos grande bloques: aquel soportado por la Comunidad y aquel soportado a nivel empresarial. En este caso se analizará sólo el soportado por la Comunidad.

Tabla 2.4.1-6 Características esenciales de Bareos

Producto/Fabricante	Garantía de copia y recuperación consistente	Garantía de continuidad de servicio y negocio	Garantía de seguridad de la información	Garantía de compatibilidad con aplicaciones de negocio	Garantía de almacenamiento	Garantía de automatización y retención de la información	Garantía de usabilidad e informes
Bareos	Bareos cumple con la garantía de copia y recuperación consistente de todos los datos críticos para el negocio y que dichos datos pueden recuperarse en, al menos, el mismo entorno donde estaban de forma consistente.	Bareos cumple con la garantía de continuidad de servicio y negocio porque el es capaz de reconstruirse o recuperarse sin perder ninguna referencia de los *backup* ya realizados.	Bareos cumple con la garantía de seguridad de la información porque es capaz de implementar algoritmos de cifrados asimétricos.	Bareos cumple parcialmente con la garantía de compatibilidad de aplicaciones de negocio. Faltan más compatibilidad con otras aplicaciones de mercado.	Bareos cumple con la garantía de almacenamiento porque es capaz de salvaguardar los datos en al menos un tipo almacenamiento (cinta, disco o nube). En este caso no se integra con cloud.	Bareos cumple con la garantía de automatización y retención de la información porque se puede automatizar tanto las copias de seguridad como las recuperaciones y proteger los datos con retención.	Bareos cumple con la garantía de usabilidad e informes porque dispone de una interfaz gráfica intuitiva y ágil. También tiene un apartado especial para informes.

2.4.2. *Software* de *backup* Propietario

Ilustración 2.4.2-1 Cuadrante Magico de Gartner de Backup

Source: Gartner (June 2016)

Commvault.com. (2017). *Enterprise Backup and Recovery | Backup & Recovery Solutions*. [online] Disponible en: https://www.commvault.com/itleaders [Accedido 29 de Marzo del 2017].

Simpana

1) Que es Simpana (https://www.commvault.com/)

Simpana es una solución de copia de seguridad comercial (su código está restringido al público) cuyo propietario es CommVault. Esta aplicación ha sufrido muchos cambios desde hace años pero, cabe destacar su estrategia de compatibilidad. Gran parte de su éxito se debe a que invierte gran parte de su presupuesto en el desarrollo de API para compatibilizar su *software* con el resto de aplicaciones y tipos de almacenamiento.

Su infraestructura está basada en tres pilares: Commserve (gestor principal de *backup*), Media agent (gestor de datos a copias o restaurar) y Disk Agent (agente de *backup* en cliente). Simpana es fácil de usar y eficiente, a la vez que

ofrece multitud de compatibilidades con aplicaciones de terceros. Cabe destacar que el destino del *backup* (storage) es indenpendiente, es decir, puede ser almacenamiento físico on-premise (disco o cinta) y también cloud como, por ejemplo, Amazon.

2) Historia

Commvault (empresa que fundo Simpana) se formó en 1988 como un grupo de desarrollo dentro de Bell Labs, y más tarde se designó como una unidad de negocio estratégica de AT & T Network Systems. En 1996, fue incorporada como una compañía independiente.

En los 20 años transcurridos desde entonces, Commvault ha experimentado un tremendo crecimiento. Ha sido pionera en numerosas innovaciones en la industria y se ha establecido como un líder de gestión de datos e información respetado. De hecho, el cuadrante mágico de Gartner para aplicaciones de copia de seguridad y recuperación de centro de datos nombró en 2016 a Commvault a Leader por sexto año consecutivo.

3) Características esenciales y comunes

Tabla 2.4.2-1 Características esenciales de Simpana

Producto/Fabricante	Garantía de copia y recuperación consistente	Garantía de continuidad de servicio y negocio	Garantía de seguridad de la información	Garantía de compatibilidad con aplicaciones de negocio	Garantía de almacenamiento	Garantía de automatización y retención de la información	Garantía de usabilidad e informes
Simpana	Simpana cumple con la garantía de copia y recuperación consistente de todos los datos críticos para el negocio y que dichos datos pueden recuperarse en, al menos, el mismo entorno donde estaban de forma consistente.	Simpana cumple con la garantía de continuidad de servicio y negocio porque el es capaz de reconstruirse o recuperarse sin perder ninguna referencia de los *backup* ya realizados. De hecho es capaz de recuperar su propio catálogo de *backup* corrompido por agentes externos a él como, por ejemplo, ante un corte de red. Además ofrece la posibilidad de recuperar la información en cloud para que el negocio continúe su actividad.	Simpana cumple con la garantía de seguridad de la información porque es capaz de implentar diferentes algoritmos de cifrados como, por ejemplo, Blowfish, AES, 3-DES…	Simpana, sin duda, cumple con la garantía de compatibilidad de aplicaciones de negocio. Desde distintos formatos de ficheros hasta el gran abanico de aplicaciones como, por ejemplo, Microsoft SQL server, Informix, DB2, Exchage, SAP,…	Simpana cumple con la garantía de almacenamiento porque es independiente del fabricante del almacenamiento ya sea almacenamiento físico o en la nube.	Simpana cumple con la garantía de automatización y retención de la información porque se puede automatizar tanto las copias de seguridad como las recuperaciones y proteger los datos con retención.	Simpana cumple con la garantía de usabilidad e informes porque dispone de una interfaz gráfica intuitiva y ágil basada en diseño web. Los informes que genera ofrecen una visión del entorno muy completa pero si esto es insuficiente, la herramienta permite generar informes personalizados.

IBM Spectrum Protect

1) Que es IBM Spectrum Protect (http://www-03.ibm.com/software/products/es/spectrum-protect-family)

   IBM Spectrum Protect es una solución de copia de seguridad comercial (su código está restringido al público) cuyo propietario es IBM. Esta aplicación ha evolucionado a lo largo de los años pero sus modificaciones no han hecho que se destaque dentro del mercado de *backup*. Quizás el aspecto a destacar sea que utiliza una instancia de DB2 como su base de datos (eliminando así las limitaciones arquitectónicas de la base de datos anterior de TSM).

   Su infraestructura está basada en cliente/servidor. No todas la aplicaciones que están presentes en el mercado están soportas por esta aplicación de *backup*. Además el destino de los *backup* no está abierto a todos los fabricantes de almacenamiento (físico o cloud).

2) Historia

   Tivoli empezó a partir de un proyecto realizado en el Almaden Research Center de IBM en 1988 para respaldar sistemas VM / CMS.

   Desde el 2015 este producto cambio su nombre pasando a ser IBM Spectrum Protect pero su funcionabilidad sigue siendo la misma. Este nuevo enfoque es una forma de renovar su imagen estancada desde 1988.

3) Características esenciales y comunes

Tabla 2.4.2-2 Características esenciales de IBM Sprectum Protect

Producto/Fabricante	Garantía de copia y recuperación consistente	Garantía de continuidad de servicio y negocio	Garantía de seguridad de la información	Garantía de compatibilidad con aplicaciones de negocio	Garantía de almacenamiento	Garantía de automatización y retención de la información	Garantía de usabilidad e informes
IBM Spectrum Protect	IBM Spectrum Protect cumple con la garantía de copia y recuperación consistente de todos los datos críticos para el negocio y que dichos datos pueden recuperarse en, al menos, el mismo entorno donde estaban de forma consistente.	IBM Spectrum Protect cumple con la garantía de continuidad de servicio y negocio porque sus datos son almacenados en una base de datos DB2 cuyo propietario también es IBM. Esto hace que tanto el *software* de *backup* como la base de datos sea de IBM incrementando su fiabilidad en caso de corrupción o error. El inconveniente en este apartado sea los exigentes requisitos *hadrware*.	IBM Spectrum Protect cumple con la garantía de seguridad de la información porque es capaz de implementar diferentes algoritmos de cifrados como, por ejemplo, AES.	IBM Spectrum Protect no es compatible con todas las aplicaciones, de hecho, para intentar garantizar esto IBM vende otros productos (suyos o de terceros) para cubrir esta carencia como, por ejemplo, para hacer *backup* de Microsoft Share Point se debe adquirir DocAve de Avepoint.	IBM Spectrum Protect cumple con la garantía de almacenamiento porque es capaz de salvarguardar los datos tanto en almacenamiento físico com o en almacenamiento virtual (cloud).	IBM Spectrum Protect cumple con la garantía de automatización y retención de la información porque se puede automatizar tanto las copias de seguridad como las recuperaciones y proteger los datos con retención	IBM Spectrum Protect cumple con la garantía de usabilidad e informes porque dispone de una interfaz gráfica intuitiva y ágil basada en diseño web. En versiones anteriores a la presente, este aspecto era uno de los puntos débiles del producto. Los informes que genera ofrecen una visión completa del entorno pero si esto es insuficiente, la herramienta permite generar informes personalizados.

EMC Networker

1) Que es EMC Networker (https://spain.emc.com/data-protection/networker.htm)

   EMC Networker es una solución de copia de seguridad comercial (su código está restringido al público) cuyo propietario es EMC. Esta aplicación ha evolucionado a lo largo de los años siendo EMC líder de almacenamiento y *backup*.

   Su infraestructura está basada en cliente/servidor y almacenamiento especializado. Es compatible con todas las aplicaciones que están presentes en el mercado y su fuerte está en el almacenamiento especializado que pone a disposición de los clientes. Dicho almacenamiento consigue los mejores resultados de deduplicación de todo el mercado.

2) Historia

   EMC Networker empezó siendo Legato NetWorker en 1988 cuyo propieratios eran cuatros personas que trabajaban juntas en Sun Microsystems: Jon Kepecs, Bob Lyon, Joe Moran and Russell Sandberg.

   Networker es absorbida por EMC cuando ésta adquiere a la empresa Legato en Octubre del 2003. A partir de entonces su producto ha ido evolucionando hasta la actualidad estando siempre entre los primeros puestos dentro del cuadrante mágico de gartner.

3) Características esenciales y comunes

Tabla 2.4.2-3 Características esenciales de EMC Networker

Producto/Fabricante	Garantía de copia y recuperación consistente	Garantía de continuidad de servicio y negocio	Garantía de seguridad de la información	Garantía de compatibilidad con aplicaciones de negocio	Garantía de almacenamiento	Garantía de automatización y retención de la información	Garantía de usabilidad e informes
EMC Networker	EMC Networker cumple con la garantía de copia y recuperación consistente de todos los datos críticos para el negocio y que dichos datos pueden recuperarse en, al menos, el mismo entorno donde estaban de forma consistente.	EMC Networker cumple con la garantía de continuidad de servicio y negocio porque la herramienta es capaz de reconstruirse o recuperarse sin perder ninguna referencia de los *backup* ya realizados.	EMC Networker cumple con la garantía de seguridad de la información porque es capaz de implementar diferentes algoritmos de cifrados como, por ejemplo, AES.	EMC Networker es compatible con todas las aplicaciones de negocio. Desde distintos formatos de ficheros hasta el gran abanico de aplicaciones como, por ejemplo, Microsoft SQL server, Informix, DB2, Exchage, SAP,…	Este es el punto fuerte de EMC Networker. El almacanemiento que ofrece EMC está espcialemente salvaguarda los datos de forma deduplicada garantizando los ratios más elevados de dedplicación en todo el mercado. Además, una vez hecho el backp se puede replicar dicho *backup* a almacenamiento virtual (cloud).	EMC Networker cumple con la garantía de automatización y retención de la información porque se puede automatizar tanto las copias de seguridad como las recuperaciones y proteger los datos con retención.	EMC Networker cumple con la garantía de usabilidad e informes porque dispone de una interfaz gráfica intuitiva y ágil basada en diseño web. Los informes que genera ofrecen una visión completa del entorno pero si esto es insuficiente, la herramienta permite generar informes personalizados

Veeam

1) Que es Veeam Backup and Replication (https://www.veeam.com/es/vm-backup-recovery-replication-software.html)

   Veeam *Backup* and Replication es una solución de copia de seguridad comercial (su código está restringido al público) cuyo propietario es Veeam. Esta aplicación es relativamente joven pero su introduccion al mercado ha revolucionado al resto de empresas que se dedican a la copia de seguridad. La fortaleza de esta aplicación son los tiempos de RTO y RPO ya que se garantizan por debajo de los 15 minutos.

   Su infreastructura está basada en cliente/servidor. La última versión de este producto permite guardar los *backups* en la nube y aumentar su escalabilidad horizontal.

2) Historia

   En 2008 nació el producto de replicación y copia de seguridad *Backup* & Replication con una plantilla de 10 empleados. Desde entonces, el producto ha ido evolucionando hasta y en 2016 se posicionó entre los 5 mejores productos de *backup*.

3) Características esenciales y comunes

Tabla 2.4.2-4 Características esenciales de Veeam

Producto/Fabricante	Garantía de copia y recuperación consistente	Garantía de continuidad de servicio y negocio	Garantía de seguridad de la información	Garantía de compatibilidad con aplicaciones de negocio	Garantía de almacenamiento	Garantía de automatización y retención de la información	Garantía de usabilidad e informes
Veeam	Veeam Backup and Replication garantiza la copia de todos los datos críticos para el negocio y que dichos datos pueden recuperarse en, al menos, el mismo entorno donde estaban de forma consistente.	Veeam Backup and Replication cumple con la garantía de continuidad de servicio y negocio porque la herramienta es capaz de reconstruirse o recuperarse sin perder ninguna referencia de los *backup* ya realizados. De hecho, es aquí donde reside su fortaleza porque garantizan un RTO y un RPO inferior a 15 minutos.	Veeam Backup and Replication cumple con la garantía de seguridad de la información porque es capaz de implementar diferentes algoritmos de cifrados como, por ejemplo, AES.	Veeam Backup and Replication es compatible con casi todas las aplicaciones de negocio pero hay un abanico muy grande donde este producto no presta servicio como, por ejemplo, toda la familia GNU Linux	Veeam Backup and Replication cumple con la garantía de almacenamiento porque es capaz de salvaguardar los datos tanto en almacenamiento físico como en almacenamiento virtual (cloud).	Veeam Backup and Replication cumple con la garantía de automatización y retención de la información porque se puede automatizar tanto las copias de seguridad como las recuperaciones y proteger los datos con retención.	Veeam Backup and Replication cumple con la garantía de usabilidad e informes porque dispone de una interfaz gráfica intuitiva y ágil basada en diseño web. Los informes que genera ofrecen una visión completa del entorno pero si esto es insuficiente, la herramienta permite generar informes personalizados.

2.4.3. Escenarios de la posible solución

Tal y como hemo visto en el apartado anterior, existe multitud de aplicaciones que se dedican a la copia de seguridad. Cada una de ellas está desarrollada para alcanzar los objetivos que demanden la empresa donde se instalará el producto. Es en este apartado donde se enumeran los posibles escenarios dependiendo del volumen de información y de capacidades de red de la empresa.

2.4.3.1. Microempresa

Las microempresas son negocios que tienen un máximo de 10 trabajadores en su plantilla. Al ser tan reducida su plantilla, el volumen de información a salva guardar no es elevado (no superaría el 500GB/año), pudiendo optar desde aplicaciones opensoruce a herramientas comerciales.

Por otro lado, las comunicaciones contratadas con proveedores de Internet suelen ser ADSL cuya capacidad (velocidad de transferencia de datos) para realizar las copias de seguridad en la nube y las restauraciones de datos es suficiente.

Dentro del análisis que se ha hecho en el apartado 2.4 podemos ver que los candidatos a cubrir este tipo de empresa podrían ser: Amanda, BackupPc, Bacula y Bareos. Los cuatro productos no implican un coste de adquisición. Por tanto, se debería de analizar los Sistemas de la Información asociados a la empresa.

Por ejemplo, si la empresa sólo se gestiona a través de ficheros ofimáticos cualquiera de los productos serviría pero si la empresa tiene una base de datos (por ejemplo, mysql) se debería elegir un producto que, o bien se integre con la base de datos; o bien ofrezca la opción de ejecutar scripts antes y después del *backup*, con el objetivo de parar la base datos, hacer *backup* a fichero de la base de datos y posteriormente arrancar la base de datos.

2.4.3.2. Empresa pequeña

La empresa pequeña se define como un negocio cuya plantilla de trabajadores oscila entre 11 y 49. La cantidad de trabajadores en este tipo de empresa implica que el volumen de información a salva guardar empieza a ser considerable (no superaría el 2TB/año). No obstante, podría optar desde aplicaciones opensoruce a herramientas comerciales.

Por otro lado, las comunicaciones contratadas con proveedores de Internet suelen ser ADSL o ADSL con fibra cuya capacidad (velocidad de transferencia de datos) para realizar las copias de seguridad en la nube y las restauraciones de datos es suficiente. En este caso, como el volumen empieza a ser considerable, tanto los tiempos de copia y recuperación como el almacenamiento en la nube se deben estudiar por motivos de costes.

Dentro del análisis que se ha hecho en el apartado 2.4 podemos ver que los candidatos a cubrir este tipo de empresa podrían ser: Amanda, BackupPc, Bacula, Bareos y Veeam. Los cuatro primeros productos no implican un coste (ver apartado 2.5.5) de adquisición pero el último sí.

2.4.3.3. Empresa mediana

La empresa mediana se define como un negocio cuya plantilla de trabajadores oscila entre 50 y 250. La cantidad de trabajadores en este tipo de empresa implica que el volumen de información a salva guardar es considerable (no superaría los 10TB/año). En este caso, posiblemente las herramientas opensource no cubran todas las necesidades y se deba analizar las aplicaciones comerciales.

Por otro lado, las comunicaciones contratadas con proveedores de Internet suelen ser ADSL con fibra cuya capacidad (velocidad de transferencia de datos) para realizar las copias de seguridad en la nube y las restauraciones de datos puede que no sea suficiente. Por ello, o bien se puede optar por descartar la solución; o bien por salvaguardar un volumen reducido de información.

Dentro del análisis que se ha hecho en el apartado 2.4 podemos ver que los candidatos a cubrir este tipo de empresa podrían ser: Amanda, Bacula, Simpana, IBM Spectrum Protect, EMC Networker y Veeam. Los cuatro primeros productos no implican un coste (ver apartado 2.5.5) de adquisición pero el resto sí.

2.4.3.4. Empresa grande

La empresa grande se define como un negocio cuya plantilla de trabajadores supera los 250 individuos. La cantidad de trabajadores en este tipo de empresa implica que el volumen de información a salva guardar es significativo (superaría los 10TB/año). En este caso, posiblemente las herramientas opensource no cubran todas las necesidades y se deba analizar las aplicaciones comerciales.

Por otro lado, las comunicaciones contratadas con proveedores de Internet suelen ser líneas simétricas de fibra cuya capacidad (velocidad de transferencia de datos) para realizar las copias de seguridad en la nube y las restauraciones de datos pueden o no ser suficientes. Por ello, o bien se puede optar por descartar la solución; o bien por salvaguardar un volumen reducido de información.

Dentro del análisis que se ha hecho en el apartado 2.4 podemos ver que los candidatos a cubrir este tipo de empresa podrían ser: Simpana, IBM Spectrum Protect, EMC Networker y Veeam. Todos productos implican un coste (ver apartado 2.5.5) de adquisición.

2.4.3.5. Costes de las aplicaciones

Con el objetivo de aportar una visión del coste que tendrá cada uno de los 4 productos comerciales y de la nube, se ha hecho comparativa de precio por TB y por año. Una puntualización, el precio es orientativo porque de cara a un entorno real donde se conjugan elementos hardware (servidores de *backup*, cabinas de almacenamiento, cables de fibra, etc.) con elementos *software* (licenciamiento de servidores Windows, Linux, deduplicación, replicación, etc.) el precio puede variar considerablemente en función de la cantidad. De hecho, los precios que se muestran en la segunda tabla sólo se ha considerado el alojamiento del almacenamiento, no el coste asociado a la transferencia de datos.

Tabla 2.4.3.5-1 Comparativa de productos comerciales con *backup on-premise*

Producto	Capacidad TB	Coste HW	Precio/TB/1año
IBM Spectrum Protect	100	52.500,00 €	525,00 €
Simpana	100	49.800,00 €	498,00 €
EMC Networker	100	48.500,00 €	485,00 €
Veeam	100	46.800,00 €	468,00 €

Tabla 2.4.3.5-2 Comparativa de productos comerciales con *backup cloud*

Producto	Capacidad	Precio/TB/1año
Azure	1	552,24 €
Amazon	1	555,73 €
CloudMe	1	1490,00€
Google	1	2146,20 €

A continuación se expondrá un ejemplo tradicional en la adquisición de una infraestructura de backup (sin incluir el coste de las licencias):

Se supone que la empresa tiene una necesidad copias de seguridad y ha estimado que se necesitan 100TB de backup para cubrir esta necesidades durante los próximos 4 años (50TB en cada CPD y la empresa tiene 2 porque su infraestructura está en alta disponibilidad).

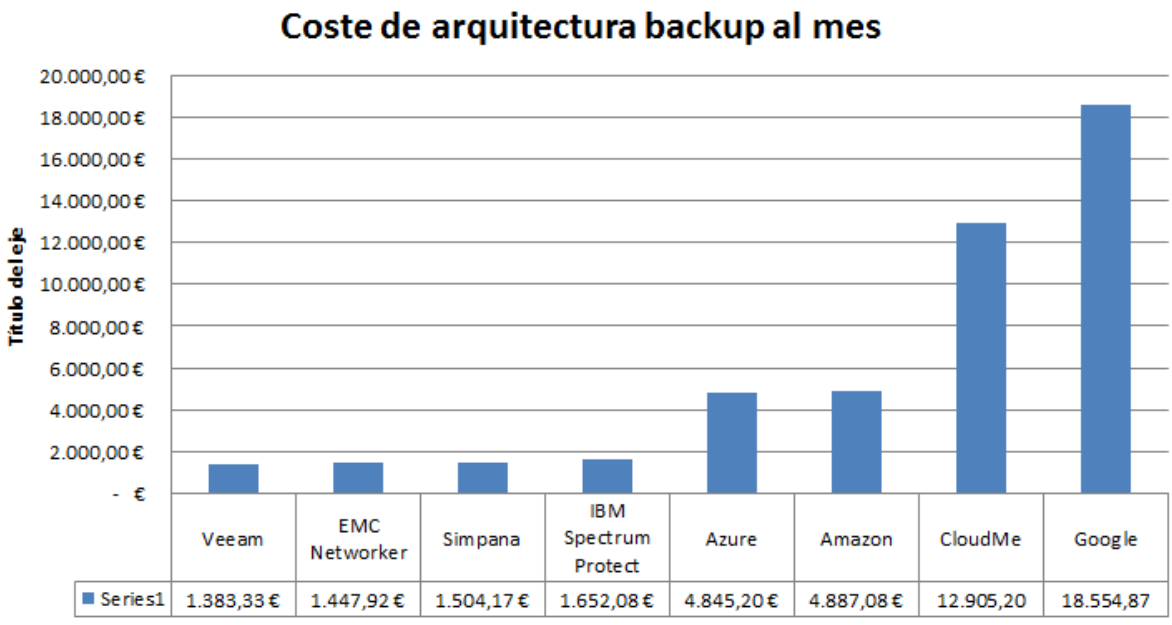
Las opciones económicas con tecnología on-premise son las siguientes:

Tecnología	Servidores	Almacenamiento	Total costes HW y soft (sin licencias)	Si recuperamos 100TB (contingencia) en un año	Coste mensual incluyendo el restorede 100TB de ese año
IBM pectrum Protect	26.800,00 €	52.500,00€	79.300,00 €	0	1.652,08 €
Simpana	22.400,00 €	49.800,00 €	72.200,00 €	0	1.504,17 €
EMC Networker	21.000,00 €	48.500,00 €	69.500,00 €	0	1.447,92 €
Veeam	19.600,00 €	46.800,00 €	66.400,00 €	0	1.383,33 €

Las opciones económicas con tecnología on-premise son las siguientes:

Tecnología	Sólo Precio/ 100TB/4años	Si recuperamos 100TB (contingencia) a 0,114€ GB en un año	Coste mensual incluyendo el restorede 100TB de ese año
Azure	220.986,00€	11.673,60€	4.845,20€
Amazon	222.292,00€	12.288,00€	4.887,08€
CloudMe	596.000,00€	23.449,60€	12.905,20€
Google	858.480,00€	32.153,60€	18.554,87€

Si comparamos ambas tecnologías y se dibujan en una gráfica que obtendrá lo siguiente.



Como conclusión, en base los TB estimados por la empresa, la infraestructura de backup basada en tencologóa on-premise es más barata que la tecnología cloud.

2.5. Requisitos físicos de la solución (alimentación eléctrica, espacio en rack, etc.)

Una vez analizado los diferentes productos y escenarios de la solución de *backup*, corresponde instalarlo en la empresa. En el caso de un escenario donde todos los servicios estén virtualizados, no requerirá de una instalación física en la empresa pero, en caso contrario, sí se será necesario.

Antes de empezar con la instalación física se debe analizar los siguientes aspectos:

- Espacio físico y accesibilidad

  El espacio físico donde se alojen los dispositivos eléctricos de *backup* es fundamental. Tanto las cabinas de discos como las librerías físicas de cintas ocupan un espacio físico dentro del habitáculo que debe ser lo suficientemente holgado como para realizar tareas de mantenimiento y/o transporte. Por ejemplo, una cabina de cintas físicas cuya tarea diaria es el intercambio de cintas (vaulting) debe tener estar ubicado en un espacio lo suficientemente grande para que un operador puede extraer e introducir las cintas.

- Electricidad

  Hasta el momento, para que funcione cualquier dispositivo de *backup* necesita conectarse eléctricamente. Dependiendo del producto, necesitará más o menos potencia eléctrica (dependiendo de sus características hardware). Es en este punto donde tanto el informático como el electricista deben analizar la potencia y el tipo de conectividad eléctrica demandada. Por ejemplo, una cabina de discos para *backup* requiere de dos tomas monofásicas de 16 Amperios. Cada toma eléctrica debe estar distribuida en una rama eléctrica diferenciada para garantizar la alta disponibilidad (esto ve en el apartado 2.6).

- Climatización

  Probablemente uno de los puntos más delicados. No sólo es necesario que habitáculo deba estar climatizado, es decir, mantener una temperatura determinada durante todo el año, sino, otros aspectos como la humedad y la polución del aire hacen que los aparatos se conserven mejor y se produzcan menos averías. En el caso de la humedad, hace que la ionización producida por el funcionamiento eléctrico de los distintos dispositivos se reduzca y, en el caso de la polución en el aire, hace que las cintas físicas de plástico y los componentes hardware de las cabinas no se deterioren con tanta rapidez.

- Conectividad de datos

  Los cables de red y de fibra es otro factor físico a tener en cuenta. No sólo porque es el elemento de comunicación entre la herramienta de *backup* con el resto de aplicaciones, sino, porque dependiendo del modelo y del número de cables, puede llegar a suponer un problema de espacio en su instalación. Por ejemplo, es espacio físico que pueden ocupar 24 cables de fibra monomodo es inferior a 24 cables de red UTP-Catergoría 6.

- Separación física entre datos y electricidad

  Aunque parezca un dato poco relevante, actualmente existen centros de procesamiento de datos donde en un mismo rail se juntan los cables eléctricos y los cables de datos (red y/o fibra). El flujo electromagnético generado por los cables de corriente eléctrica modifica la señal de los cables de comunicación llegado a provocar cortes de comunicación e incluso inconsistencia en los datos transmitidos. Es por ello muy importante separar ambos elementos físicos a una determinada distancia.

2.6. Análisis, diseño y construcción de infraestructura de *backup*

Una vez instalado el hardware (hay modelos en los que no hace falta) se debe analizar dos puntos de vista:

- **Como será la infraestructura de *backup***: este análisis a su ve se dividirá en do partes: físico y lógico.
    - En el caso físico, se debe detallar la ubicación física de los servidores, la conectividad eléctrica y de los datos (red y fibra) tal y como se ha visto en el apartado anterior.
    - En el caso lógico, se debe detallar qué unidades de negocio cubre dicha solución, por ejemplo, la herramienta de *backup* contiene un agente de integración de SQL para la base de datos core de negocio.

- **Qué objetivos se deben alcanzar a través de análisis, diseño y construcción de la infraestructura de *backup***. Ente estos objetivos cabe destacar la garantía de continuidad del negocio y del servicio, la garantía de la alta disponibilidad del servicio de *backup* y del dato de negocio, la protección del dato (LOPD), el control de acceso a la herramienta de *backup*, la integración de la nueva infraestructura de *backup* en la empresa y la garantía de la máxima eficiencia del *software* de *backup*. A continuación, se explica cada uno de los objetivos:

2.6.1. Garantizar la continuidad del negocio y del servicio

Estos dos objetivos se deben aplicar tanto a la herramienta de *backup* como las aplicaciones y bases de datos de la empresa.

Continuidad del negocio

Toda aquella aplicación y/o base de datos que genere ingresos o rentabilidad debe estar identificada y clasificada según su importancia para el negocio. Existe un proceso de análisis dentro de la empresa llamado BIA (*Business Impact Analysis*) cuyo resultado es un informe donde se especifican aquellas aplicaciones/bases de datos más críticas para negocio y, en caso de desastre, accidente o emergencia, pueden suponer desde la parada de la actividad empresarial hasta el cierre del negocio.

Continuidad del servicio

Tiene dos objetivos fundamentarles de recuperación: tiempo máximo en el que se debe alcanzar un nivel de servicio mínimo tras una pérdida del servicio (RTO) y fecha y hora desde la que se restaurarán los datos tras una pérdida del servicio (RPO). Además, factores como el ancho de banda en las líneas de comunicación, velocidad del disco y agilidad en el proceso de recuperación son relevantes de cara a la continuidad del negocio.

2.6.2. Garantizar la alta disponibilidad del servicio de *backup* y del dato de negocio.

Alta disponibilidad del servicio de *backup*

Se debería diseñar una solución que esté en alta disponibilidad, es decir, en caso de que el nodo o servidor que contenga el *software* de *backup* falle, este servicio se debería balancear o conmutar a otro nodo o servidor para seguir ofreciendo el servicio de copia de seguridad.

Alta disponibilidad del dato de negocio

Se debería diseñar una solución donde el dato copiado deba estar accesible en cualquier momento. Para ello es muy interesante la replicación del dato copiado entre los diferentes tipos de almacenamiento del mercado. De tal modo que, en caso de que falle alguno ellos, siempre se pueda acceder desde el otro.

2.6.3. Proteger el dato copiado: LOPD

La Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal debe ser analizada e integrada en la infraestructura de *backup*. No todas las copias de seguridad contienen datos sensibles como, por ejemplo, datos personales, las libertades públicas y los derechos fundamentales de las personas físicas, y especialmente de su honor, intimidad y privacidad personal y familiar, pero, aquellos que sí lo contienen deben tratarse de forma especial.

La empresa debe facilitar información sobre el tipo del dato que se necesita copiar (¡ojo! no su contenido). Dependiendo la naturaleza del dato, el nivel de seguridad LOPD se puede dividir en:

- **Nivel bajo**: si lo datos contienen datos identificativos, características personales, circunstancias sociales, datos profesionales, empleo, información comercial, datos económicos-financieros y de seguros y datos de transacciones.

- **Nivel medio**: si lo datos contienen infracciones administrativas o penales, datos de administraciones tributarias, entidades financieras, entidades gestoras y servicio comunes de seguridad social, mutuas de accidentes de trabajo y enfermedades profesionales, datos de definición de la personalidad o del comportamiento.

- **Nivel alto**: ideología, afiliación sindical, religión, creencias, origen racial, salud, vida sexual. Datos recabados con fines policiales. Datos derivados de actos de violencia de género.

Ante el robo de datos de *backup* y su posterior publicación, la empresa podría ser fuertemente sancionada con una multa económica por no proteger adecuadamente los datos salvaguardados. Desde el punto de vista de *backup*, se debe solicitar a la empresa información sobre cuáles son y dónde están dichos datos y, posteriormente, aplicar el tipo de cifrado más adecuado en función del nivel de seguridad de la LOPD.

2.6.4. Limitar el acceso a la herramienta de *backup*

Dentro de ámbito empresarial donde el número de personas con perfiles informáticos es elevado, es muy aconsejable definir un conjunto de perfiles con el objetivo de reducir posibles riesgos asociados a errores humanos.

En las empresas, generalmente medianas y grandes, el departamento informático de sistemas suele estar jerarquizado en varios niveles: director, administrador, técnicos y/o operadores. Todos estos niveles tienen su responsabilidad pero puede existir un punto de "conflicto" entre todos ellos a la hora de gestionar una herramienta, en este caso, la herramienta de copias de seguridad.

Para entender mejor el "conflicto" veamos un ejemplo. Supongamos que el director de sistemas se equivoca y borra un registro de una tabla de una base de datos SQL core para el negocio. El director, con la mejor voluntad, intentará recuperar dicha información accediendo al *software* de *backup* pero, al lanzar la recuperación se equivoca y recupera todos los datos de hace un mes… ¡Catástrofe!

Si el usuario del director sólo tuviera los permisos adecuados, es decir, sólo permisos para ver si existe una copia de la base de datos, éste hubiera mandado un correo a los operadores o al administrador para efectuar la recuperación de forma correcta. Aquí es donde toma importancia la definición de perfiles de los usuarios de la aplicación de *backup*.

2.6.5. Integrar la nueva infraestructura de *backup* en la empresa

La integración de nueva infraestructura de *backup* en la empresa debería ser lo menos disruptiva posible para que el proyecto no genere negatividad. Si se cumple este aspecto, favorecerá las futuras actualizaciones o modificaciones del propio producto sin que haya oposición al cambio (aunque siempre lo habrá).

A colación de lo anterior, la nueva infraestructura de *backup* se debe adaptar a la empresa y viceversa. Extendiendo la anterior idea, el *software* de *backup* debería cubrir las necesidades de copia y restauración de seguridad de la empresa, en el menor tiempo posible y de la forma más eficiente. Conjuntamente, la empresa debería facilitar información de todos los sistemas que se tienen hacer copia, así como, de una franja horaria donde se garantice que el dato que se esté copiando no se esté modificando. En consecuencia, la empresa debería tener muy claro cuando se ejecutan los procesos críticos de negocio, qué datos modifican y cuánto tiempo tardan.

Por otro lado, dependiendo del *software* de *backup* adquirido y del tipo de aplicaciones que tenga instalada la empresa, puede que no se cubran todas las necesidades exigidas. Para confirmar esto, se debe analizar todos los sistemas presentes y futuros (en la medida de lo posible) y, en caso de identificar alguna incompatibilidad, intentar proponer soluciones alternativas.

2.6.6. Garantizar la máxima eficiencia del *software* de *backup*

De poco sirve tener el mejor *software* de *backup* del mercado si no se ha analizado y diseñado una infraestructura que garantice la eficiencia de la copia de seguridad en la empresa.

Según se describió en el punto 2.4.3, existen diferentes escenarios a partir de los cuales una empresa se puede clasificar. Uno de los factores fundamentales para esta clasificación es el tipo de comunicación que tiene la empresa. En consecuencia, se podrían dar los siguientes casos:

- **Empresa con gran ancho de banda en la comunicación tanto interna como externa**

  Este modelo de empresa tendrían accesos internos basados en tecnología de comunicación cuya velocidad de navegación sea superior a los 10Gbps. Ejemplo de este tipo de tecnología serían los canales de fibra o FCoE. Además, accesos externos estarían basados en tecnología de comunicación cuya velocidad de navegación sea superior a los 4Gbps. Ejemplo de este tipo de tecnología serían la fibra oscura de Telefónica.

- **Empresa con gran ancho de banda en la comunicación interna pero con limitaciones en la comunicación externa**

  Este modelo de empresa tendrían accesos internos basados en tecnología de comunicación cuya velocidad de navegación sea superior a los 10Gbps pero, los accesos externos estarían basados en tecnología de comunicación cuya velocidad de navegación trabaje a Mbps. Ejemplo de este tipo de tecnología sería DSL simétrico.

- **Empresa con limitación de ancho de banda en la comunicación tanto interna como externa.**

  Este modelo de empresa tendría accesos internos basados en tecnología de comunicación cuya velocidad de navegación trabaje a Mbps. Así mismo, los accesos externos estarían basados en tecnología de comunicación cuya velocidad de navegación trabaje a Mbps.

2.6.7. Almacenamiento de *backup*: cinta, disco o nube

Dependiendo del *software* de *backup*, se tendrá la oportunidad de elegir o descartar uno o varios modelos de almacenamiento de datos. En cualquiera de los casos, es importante señalar las ventajas y desventajas de cada uno de los modelos.

Almacenamiento en cinta

Una cinta de *backup* es un dispositivo físico cuyo funcionamiento se basa en la traslación de una cinta magnética mientras se copia o restaura el dato. Hasta el momento, sólo existen dos métodos de escritura en este tipo de dispositivo: *start-stop* o *streamming*.

- **El modo de escritura *star-stop*** donde la cinta debe acelerarse para comenzar a escribir y parar cuando termina. Entre el inicio de la primera escritura y el siguiente, aparecen huecos donde no se aprovecha el espacio.

- **El modo de escritura *streaming*** consiste en escribir de forma continua sin parar. De este modo no se dejan espacios en blanco y se aprovecha todo el espacio de la cinta.

Las características que se pueden encontrar en este tipo de tecnología son:

- **Compresión y capacidad**: la tecnología LTO Ultrium permite, dependiendo de dato copiado y de la generación de la cinta, comprimirlo alcanzando capacidades de hasta 120TB en un sola cinta (generación 10 LTO Ultrium).

- **WORM**: funcionalidad desarrollada desde la generación 3 para garantizar que los datos copiados de una empresa y almacenados en un cinta con esta tecnología no sean alterados. Muy útil frene a procedimientos legales y reglamentarios.

- **Encriptación**: al ser un dispositivo hardware, el cifrado de datos copiados es útil en caso de robo de dicho dispositivo.

- **Velocidad de escritura**: la velocidad de escritura en la última generación de cinta puede llegar a alcanzar los 2750 MB/sg

- **Particionamiento**: esta característica apareció en la generación 6 de LTO Ultrium y separa la cinta en dos particiones: la primera contendrá un índice con el contenido de la cinta y la segunda será el propio contenido de la cinta.

Aunque en la actualidad se está desechando esta tecnología por alguna de las características anteriormente mencionadas pero su evolución es imparable e innegable (ver Anexo I). Cabe destacar que este tipo de tecnología es muy fiable en aspectos legales como retención de datos durante años porque las cintas son más duraderas, es rentable debido a su larga vida, es situable en el espacio, es decir, se sabe exactamente donde se ubican los datos de la empresa.

Almacenamiento en disco

La definición de un disco ha variado desde los últimos años. Lo que un principio era y es un conjunto de platos metálicos circulares que giran a un número determinado de revoluciones por minuto (HDD), ahora puede llegar un módulo de memoria no volátil (SSD).

La forma de leer y escribir en este tipo de hardware varía según su modelo: en el modelo tradicional (eje de discos) el acceso de lectura o escritura es aleatorio, es decir, el plato gira de forma continua y puede darle acceso a cualquier sector de la superficie en una fracción de segundo. Sin embargo, el modelo electrónico SSD, escribe del mismo modo que se escribe en la memoria física de un ordenador.

La velocidad de lectura y escritura de los discos, el crecimiento horizontal y vertical y la facilidad de replicación de los datos son características que favorecen a este tipo de tecnología. Por el contrario, su duración y el coste son desventajas a tener en cuenta.

Almacenamiento en nube

El almacenamiento de *backup* en la nube no se pude describir con exactitud porque es un tipo de tecnología transparente para el cliente o usuario. Esta característica es una ventaja porque despreocupa sobre problemas asociados al hardware pero, a su vez, la falta de información puede preocupar al usuario por no tener la certeza de que sus datos se están guardando de forma cifrada y en el mismo país donde se encuentra la empresa (aplicación de la LOPD).

A día de hoy, el almacenamiento en la nube tiene tasas de transferencia de datos muy bajas (poca velocidad) y, los costes asociados a salvaguardar la información, la subida y/o bajada de información, puede encarecer significativamente el coste del servicio de *backup*.

2.7. Instalación y parametrización del *software* de *backup*

Una vez analizado, diseñado y construido el entorno de *backup* se debe pensar en la instalación y parametrización del *software* de *backup*. Dependiendo de producto comprado, puede resultar más o menos sencillo pero tiene como objetivo fundamental aprovechar de la forma más eficiente todas las características del *software*.

Para rentabilizar lo máximo posible la eficiencia se abordará otro aspecto fundamental dentro de un entorno de *backup* llamado **políticas de seguridad**. Dichas políticas se basa en el dónde, cuánto y cómo debemos tratar los datos copiados. A continuación hablaremos de cada una de ellas.

- Dónde…

  ¿Dónde copio y/o recupero los datos? Es un factor a tener en cuenta dependiendo de modelo elegido. En el caso de tener un modelo on-premise, lo más eficiente es que los datos de dicho CPD sean copiados localmente y trasladados a otro centro de datos. Esto se puede hacer físicamente (vaulting) o mediante replicación de datos a otro CPD o a la nube.

  En el caso de tener un modelo basado en la nube, lo más eficiente es que los datos ubicados en la nube sean copiados con sus propias herramienta pero no trasladado de ubicación. Esto debe analizarse con detenimiento porque la legislación que rige en distintos paise y/o continentes no es la misma, por ejemplo, la LOPD de Europa no es la misma que Estados Unidos.

  En relación al anterior párrafo, cabe destacar la diferencia que existe entre dato de backup, dato de negocio y dato historificado.

  Dato de negocio

  - Cuya información está activa y disponible para cualquier usuario o cliente autorizado.
  - Puede ser o no modificado en un periodo de tiempo determinado o indeterminado.
  - La corrupción de este fichero puede suponer un problema para la continuidad del servicio o del negocio.
  - Este tipo de dato ocupa espacio en almacenamiento especializado para la actividad empresarial, no debería ocupar espacio en almacenamiento de backup.

Dato historificado

- El acceso a este dato puede ser circunstancial o bajo demanda pero siempre debe estar activo y disponible para determinados usuarios.
- No puede ser modificado porque ha cumplido con su ciclo de vida y está bajo políticas restrictivas de acceso al dato.
- La corrupción de este dato puede suponer un problema para la continuidad del servicio o del negocio.
- Este tipo de dato debería ocupar espacio en almacenamiento especializado para la historificación, no debería ocupar espacio en almacenamiento de backup ni de negocio.

Dato de backup

- Cuya información está activa y disponible para la recuperación de algún dato de negocio o de historificación.
- El acceso al dato de backup está restringido a los usuarios que deban efectuar labores de recuperación.
- La corrupción de este dato puede suponer o no un problema para la continuidad del servicio o del negocio.

- Cuánto…

¿Cuántos días retengo mis copias? El informe BIA aportará las claves para tomar esta decisión pero es el consumo de espacio en almacenamiento quien termine de contestar esta pregunta. Por ejemplo, si se está haciendo un *backup* diario de una base de datos Informix de 5TB con una retención de 15 días, la empresa o la nube deberán tener una capacidad de 75TB. Esto se traduce en un aumento de costes para la empresa porque el almacenamiento de datos es limitado (los datos siguen creciendo día a día y es algo imparable).

Otro aspecto a tener en cuenta son los ciclos de vida el dato, es decir, cuanto tiempo necesito retener mi información y que además sea válida en caso de recuperación. Imaginemos una empresa cuyo negocio trata datos médicos y no médicos. Para los primeros, se debe garantizar su permanencia durante 5 años según la Agencia Española de Protección del Dato, mientras que, la segunda, Negocio establecerá su retención.

- Cómo…

  ¿Cómo se deben ejecutar mis *backups*? Este punto está relacionado con el anterior. Sólo se mencionará que existen la deduplicación y técnicas de cómo realizar una copia de seguridad de la forma más eficiente porque en el apartado 2.8 se explicará de forma más detallada dicha información.

  Existen distintos tipos de *backup* como pueden ser *backup* completos (full *backup*), *backup* incrementales, diferenciales y de logs. Pero no hace mucho tiempo, ha surgido dos nuevos tipos de backup: *backup* sintético y *backup* incremental. Estos nuevos modelos de *backup* no sólo ahorran espacio y tiempo, sino, costes para la empresa.

Un ejemplo de políticas de seguridad podría ser el siguiente. Las políticas de *backup* que se describen en este ejemplo se basan en dos aspectos: el primero, se basa en la retención de la información copiada y, el segundo, se basa en la criticidad del proceso o servicio de negocio.

Política Oro

- ***Backup* diario**: cuya frecuencia de ejecución son todos los días menos cuando se tenga que ejecutar el *backup* semana, mensual o anual. La retención de los datos copiados sería 30 días y el tipo de *backup* sería incremental.

- ***Backup* semanal:** cuya frecuencia de ejecución es una vez a la semana, por ejemplo, todos los sábados, excepto cuando se tenga que ejecutar el *backup* mensual o anual. La retención de los datos copiados sería 12 semanas y el tipo de *backup* sería completo.

- ***Backup* mensual**: cuya frecuencia de ejecución es una vez al mes, por ejemplo, el primer domingo de cada mes excepto cuando se tenga que ejecutar el *backup* anual. La retención de los datos copiados sería 12 meses y el tipo de *backup* sería completo.

- ***Backup* anual**: cuya frecuencia de ejecución es una vez al año, por ejemplo, el tercer día de cada año. La retención de los datos copiados sería 5 años meses.

Política Plata

- ***Backup* diario**: cuya frecuencia de ejecución son todos los días menos cuando se tenga que ejecutar el *backup* semana, mensual o anual. La retención de los datos copiados sería 15 días y el tipo de *backup* sería incremental.

- ***Backup* semanal**: cuya frecuencia de ejecución es una vez a la semana, por ejemplo, todos los sábados, excepto cuando se tenga que ejecutar el *backup* mensual o anual. La retención de los datos copiados sería 6 semanas y el tipo de *backup* sería completo.

- ***Backup* mensual:** cuya frecuencia de ejecución es una vez al mes, por ejemplo, el primer domingo de cada mes excepto cuando se tenga que ejecutar el *backup* anual. La retención de los datos copiados sería 6 meses y el tipo de *backup* sería completo.

- ***Backup* anual**: cuya frecuencia de ejecución es una vez al año, por ejemplo, el tercer día de cada año. La retención de los datos copiados sería 2 años meses.

Política Bronce

- ***Backup* diario**: cuya frecuencia de ejecución son todos los días menos cuando se tenga que ejecutar el *backup* semana, mensual o anual. La retención de los datos copiados sería 7 días y el tipo de *backup* sería incremental.

- ***Backup* semanal**: cuya frecuencia de ejecución es una vez a la semana, por ejemplo, todos los sábados, excepto cuando se tenga que ejecutar el *backup* mensual o anual. La retención de los datos copiados sería 2 semanas y el tipo de *backup* sería completo.

- ***Backup* mensual**: cuya frecuencia de ejecución es una vez al mes, por ejemplo, el primer domingo de cada mes excepto cuando se tenga que ejecutar el *backup* anual. La retención de los datos copiados sería 2 meses y el tipo de *backup* sería completo.

- ***Backup* anual**: no tendría.

2.8. Pruebas y análisis de *software* de *backup* (*backup* y *restore*)

Aunque es obvio, de poco sirve una herramienta de copias de seguridad si la recuperación falla, este razonamiento es el que nos impulsa obligatoriamente llevar a cabo las pruebas de *backup* y *restore*, y analizar los resultados obtenidos.

Para saber en qué orden empezar las pruebas, se debería utilizar el documento BIA (se explicó en el 2.6 de este documento) porque mostrará los proceso más críticos para negocio. Para llevar a cabo esto, es necesario tener un entorno lo más parecido al entorno actual de la empresa donde se pueda reproducir posibles contingencias. Este entorno puede estar o no aislado pero es fundamental que dichas pruebas no afecten ni al servicio ni al negocio.

Por otro lado, las pruebas que deberían efectuarse se pueden agrupar en dos grandes bloques: pruebas de copias de datos y pruebas de recuperación de datos y/o servicios. A continuación, se explica en detalle cada una de ellas.

2.8.1. Pruebas de copias de datos

En los últimos años, la copia de seguridad ha dado un giro donde se deja de hablar de ficheros y toma mayor importancia los bloques de datos (división física de un fichero).

Recordemos que las pruebas de copias de seguridad deben garantizar que el dato copiado es consistente para su futura recuperación. Las copias de seguridad se pueden dividir en distintos tipos.

Según el tipo de *backup*

- **Completas**: tipo de copia de seguridad donde se copia todos los bloques de datos del origen y se guardan en el destino. Este tipo de copia de seguridad tiene las siguientes ventajas:
    - En una sola operación se dispone de todos los datos.
    - Mejora el RTO porque de una operación se puede restaurar los datos en el menor tiempo
    - En aquellas tecnologías donde la deduplicación se sitúa en el destino, es decir, en la cabina de discos o la nube, permite conseguir mejores ratios de deduplicación siempre y cuando el dato ni varíe significativamente ni se cifre.

Por el contrario, si no se dispone de una tecnología que no deduplica, como por ejemplo la cinta, la empresa requerirá más inversión en almacenamiento. Además, el tiempo de ejecución de este tipo de *backup* suele ser considerable, pudiendo colisionar con la ejecución de otros *backups* y consumir demasiados recursos. A modo ilustrativo ver anexo II.

- **Incrementales**: tipo de copia de seguridad donde se copia aquellos bloques que han sido modificados con respecto al *backup* completo o al *backup* incremental anterior. Cada conjunto de bloques se agrupa independientemente, es decir, no se acumula todos los bloques modificados en una misma agrupación. Este tipo de copia de seguridad tiene las siguientes ventajas:

    - Dependiendo del sistema, el espacio ocupado por este *backup* suele ser inferior con respecto a la copia completa.

    - A colación de la anterior ventaja, el tiempo de ejecución es menor que la copia completa.

    Por el contrario, este tipo de copias no favorecen mucho a la deduplicación porque sólo trasladan el dato modificado. Además, suele empeorar el RTO porque para recuperar la última copia se debe primero restaurar el *backup* completo (dependencia obligatoria) y, posteriormente, todos los *backups* incrementales hasta el momento elegido (dependencias obligatorias). A modo ilustrativo ver anexo III.

- **Diferenciales**: tipo de copia de seguridad donde se copia aquellos bloques que han sido modificados con respecto al *backup* completo o al *backup* diferencial anterior. Cada conjunto de bloques modificado se agrupa y se acumula en una misma agrupación. Este tipo de copia de seguridad tiene las siguientes ventajas:

    - Dependiendo del sistema, el espacio ocupado por este *backup* suele ser inferior con respecto a la copia completa.

    - A colación de la anterior ventaja, el tiempo de ejecución es menor que la copia completa.

    Por el contrario, este tipo de copias no favorecen mucho a la deduplicación porque sólo trasladan el dato modificado. Además, contra más veces se ejecute este tipo de *backup*, más aumentará el espacio demandando porque acumulará los cambios en una sola agrupación. Al igual que el incremental, suele empeorar el RTO porque para recuperar el último *backup* primero se debe restaurar el *backup* completo (dependencia obligatoria) y, posteriormente, el *backup* diferencial que en el caso de que su volumen de información sea elevado tardará más tiempo en recuperar (dependencias obligatoria). A modo ilustrativo ver anexo IV

- **Completo-sintético (*Synthetic Full Backup*)**: tipo de copia de seguridad donde se fusiona el mapa de bloque de datos ya copiados en uno actual y completo. En el caso de que se estén haciendo *backups* diferenciales o incrementales habrá un momento donde la empresa quiera tener de nuevo un *backup* completo con los datos guardados en *backups* anteriores. Es aquí donde el *backup* sintético aparece porque este tipo de tecnología es capaz de unir la información ya salvaguarda en un solo *backup* sintético y completo. Este tipo de copia de seguridad tiene las siguientes ventajas:

    - En una sola operación se dispone de todos los datos ya salvaguardados.

    - Consume menos recursos porque trabaja directamente con el almacenamiento de *backup*.

    - Mejora el RTO porque de una operación se puede restaurar los datos en el menor tiempo.

    A modo ilustrativo ver anexo V

- **Incrementa para siempre (*Incremental Forever*)**: tipo de copia de seguridad donde el mapa de bloque de datos es siempre el más actual y sólo se guarda aquel bloque dato modificado. A cada bloque de dato modificado se hace referencia desde el primer *backup*, por tanto, sólo se guarda el bloque modificado en cada copia (incremental para siempre). Este tipo de copia de seguridad tiene las siguientes ventajas:

    - Ahorro de espacio en almacenamiento.

    - Reducción de tiempo de ejecución.

    - Mejora el RTO porque de una operación se puede restaurar los datos en el menor tiempo

    A modo ilustrativo ver anexo VI

Según el estado del proceso de negocio

Para hacer determinadas copias de seguridad, a veces se deben parar los procesos de negocio durante un determinado tiempo porque o bien la herramienta de *backup* no incluye su compatibilidad con el *software* copiado, o bien, el *software* no está diseñado para hacer copias de seguridad mientras está en funcionamiento. Es por ello que se deben distinguir dos tipos de copias:

- **Copias en caliente**: donde la aplicación, fichero o base de datos está activo. En este caso en muy importante que la copia sea consistente (a pesar de estar activa) para garantizar la integridad del dato.

- **Copias en frío**: donde la aplicación, fichero o base de datos está parada. En este caso en muy importante que la velocidad copia sea lo mayor posible para reducir en tiempo de no disponibilidad del servicio.

2.8.2. Pruebas de recuperación de datos

Las pruebas de recuperación son fundamentales de cara a la continuidad del servicio y del negocio. Es por ello compensaría tener un entorno pruebas lo más parecido al entorno productivo de la empresa para poder realizar estimaciones equiparables.

El conjunto de pruebas de recuperación de datos se pueden dividir en:

- **Restauración completa**: aquella donde se requiere recuperar toda la información cuando se ejecutó la copia de seguridad.

- **Restauración incremental**: aquella donde sólo se requiere la recuperación parcial de aquellos datos que han sido modificados en el tiempo.

- **Restauración hasta un punto en el tiempo**: ante un desastre se requiere la recuperación de los datos hasta un momento determinado.

Otro tipo de pruebas serían aquellas que están orientadas a la alta disponibilidad del servicio de *backup* donde, en caso de fallo de uno de los nodos de *backup*, el otro nodo se activaría automáticamente y continuaría dando servicio. Este tipo de pruebas se pueden dividir en 3 apartados:

a) Pruebas de corte de corriente

Este tipo de prueba es la más agresiva porque consistiría en desenchufar eléctricamente el dispositivo de *backup* que, dependiendo de su naturaleza, se podrá realizar o no. Un ejemplo en el que no se podría realizar esta prueba sería en el caso de un *software* de *backup* basado la nube. Apagar un dispositivo nfísico que está prestando servicio a multitud de clientes supodría un grave problema para la empresa proveedora.

b) Pruebas de corte de comunicaciones

Dentro del entorno de las comunicaciones se pueden dividir en dos clases: aquellas que utilizan tecnología Ethernet y aquellas que utilizan tecnología de fibra.

a. Pruebas de corte de red

Este tipo de prueba consistiría en desconectar los cables red del servidor de *backup* para dejarlo incomunicado. Otra prueba asociada a la red sería quitar la comunicación de red al almacenamiento asociado a este entorno de *backup*.

b. Pruebas de corte de SAN

Este tipo de prueba consistiría en desconectar los cables de fibra del servidor de *backup* para dejarlo incomunicado. Otra prueba asociada a la red sería quitar la comunicación de fibra al almacenamiento asociado a este entorno de *backup*.

Ambas pruebas servirían para saber cómo se comportaría la infraestructura de *backup* en caso de corte de comunicaciones de red y de fibra que, en caso de disponer de alta disponibilidad, nunca se cortaría el servicio de *backup*.

2.8.3. La deduplicación

A lo largo de los años, las copias de seguridad han estado copiando datos sin tener en cuenta que hay partes de los propios datos idénticas a otros datos ya copiados (en el mismo backup o en otros). Dicho de otra forma, la probabilidad de que en un mismo almacenamiento haya copias con datos repetidos es elevada a pesar de utilizar técnicas como backup incrementales, diferencial, etc.

La deduplicación es una técnica que elimina el dato duplicado. Para llevar a cabo esto, las tareas que realiza las divide en 4 etapas:

1) Recoge el dato a ser salvaguardo.
2) Divide el dato en pequeños bloques de datos (esto varía según el fabricante de software o hardware).
3) Obtiene un hash de cada uno de los bloques de datos.
4) Compara el hash obtenido con todos los hash del almacenamiento de la base de datos.
    a. En caso de que el hash no haya sido encontrado, se copia el bloque de dato al almacenamiento.
    b. En caso de que el hash sí ha haya sido encontrado, se crea una referencia (link) al dato del almacenamiento y esta referencia se guarda en la copia de seguridad.

Antes de activar esta técnica se debería tener claro que, dependiendo de donde y cuando se ejecute se obtendrán ventajas e inconvenientes. Por ello, las formas de ejecutar la deduplicación se pueden dividir:

a) En referencia a dónde se ejecutan

a. Deduplicación en origen

En este caso, el trabajo lo realiza el cliente o servidor que contiene los datos a copiar. La ventaja de esta técnica es la reducción de los datos a copiar, la reducción del consumo del canal de comunicación. Sin embrago, como desventaja, el cliente o servidor deberá realizar todas operaciones y, por tanto, exigirá más ciclos de computación (consumo de procesador).

b. Deduplicación en destino

En este caso, el trabajo lo realiza almacenamiento donde están los datos copiados. La ventaja de esta técnica es la reducción de los datos a copiar y el ahorro de consumo de cpu del cliente o servidor. Sin embrago, como desventaja, el almacenamiento deberá realizar todas operaciones y, por

tanto, exigirá más ciclos de computación (consumo de procesador). Este tipo de deduplicación a su vez puede dividirse:

c. Deduplicación en origen y destino

En este caso, servidor y almacenamiento trabajan con los datos. La ventaja de esta técnica es una doble deduplicación, es decir, el origen deduplica en dato y el destino deduplica el dato ya deduplicado. La ventaja es que aumenta el ahorro de espacio pero la desventaja es que el consumo de procesador se eleva en ambos extremos.

b) En referencia a cuando se ejecutan

a. Deduplicación *Inline*:

El proceso se ejecuta al mismo tiempo que el backup y el dato se analiza antes de escribirlo a disco. Esto hace la escritura a disco más lenta pero el consumo de espacio en el almacenamiento sea mínimo.

b. Deduplicación *Offline*:

El proceso se ejecuta después del backup. Este proceso no perjudica a la velocidad de escritura pero, tiene como principal inconveniente, tener tanto espacio en disco como el dato sin deduplicar. Además, se requiere de una ventana adicional de trabajo (housekeeping) para eliminar los datos deduplicados.

Dentro del conjunto de aplicaciones de backup analizadas en el punto 2.4, todas aquellas que son de tipo opensource no disponen de esta utilidad, mientras que, todas las aplicaciones comerciales sí trabajan con la deduplicación.

3. Conclusiones

Las lecciones aprendidas en esta memoria son:

- Antes de afrontar una implantación de *software* de *backup* se debe tener claro que es un reto complejo y fundamental para la continuidad del servicio y del negocio. Dicho de otra forma, si una empresa carece de este servicio se arriesga a sufrir el cierre completo de su actividad empresarial en un momento determinado.

- Desde el inicio, la empresa debe estar implicada en el proyecto porque hay determinadas decisiones que no son responsabilidad ni del administrador de *backup* ni del proveedor que apoye o implante esta solución, como por ejemplo, la retención de los datos, los niveles de servicio (SLA), los RTO y RPO.

- Este proyecto representa una garantía de futuro para la actividad de la empresa por lo que acortar la duración de la implantación, no analizar los riesgos de forma detallada y reducir costes puede suponer un grave problema para la empresa porque perjudicarían aspectos clave para el desarrollo del proyecto.

- No se debe olvidar que, como herramienta fundamental para la empresa, debe estar dotada de alta disponibilidad a tres niveles: físico, donde debe tener redundancia eléctrica y de conectividad de datos; lógica, donde el dato del negocio debe estar accesible aunque se produzcan fallos y, por último, a nivel de servicio, donde el servicio *backup* siempre debe estar activo.

- Durante los últimos años el sector de copias de seguridad ha evolucionado en dos trayectorias diferentes: la primera se centra en las copias de seguridad ya que se pueden guardar en el almacenamiento de la nube; la segunda se centra el concepto de bloque de datos, desapareciendo los ficheros como unidad de copia. Ambas trayectorias tienen en común la deduplicación que no sólo mejora el rendimiento de las copias de seguridad, sino que ahorra costes para la empresa.

- Los cálculos que dictaminan el espacio que deberá tener el futuro almacenamiento para salvaguardar la información son cada vez más complejos porque el dato crece día a día y a esto se suma las nuevas técnicas de *backup* como la copia completa sintética o incremental para siempre cuya característica principal es reducir la información a salvaguardar.

- La ubicación del dato o el tipo de dato tiene implicaciones legales referentes a la protección de los datos (LOPD). En el primer caso, cada país y/o continente se rige por sus propias leyes por lo que su confidencialidad puede estar comprometida; mientras que en el segundo caso, la pérdida de información por no estar bien establecida su retención tiene consecuencias legales.

En base a las anteriores reflexiones se ha logrado los objetivos planteados inicialmente. Este trabajo ha seguido la planificación definida al principio de la memoria. La metodología prevista parece ser la más adecuada por la naturaleza del proyecto y no ha hecho falta introducir cambios significativos en el proyecto que garanticen el éxito del trabajo.

Las líneas de trabajo futuro que no sean podido explotar y han quedado pendientes serían aquellas donde las aplicaciones "*cobran vida*", es decir, tener un entorno real (para ello se requiere de presupuesto y tiempo) donde todos los elementos *software* y hardware estuvieran al alcance para demostrar empíricamente las ventajas y desventajas de los productos de *backup*.

4. Glosario

Backup término inglés que hace referencia a la copia de seguridad de datos.

Restore término inglés que hace referencia a la recuperación de datos.

SLA término en inglés que hace referencia a los acuerdos de nivel de servicio.

RTO término en inglés que hace referencia al objetivo de recuperación en el tiempo.

RPO término en inglés que hace referencia al objetivo de recuperación en un punto.

BIA término en inglés que hace referencia al informe llamado Análisis de Impacto de Servicio.

Deduplicacion técnica donde se eliminan los datos duplicados.

LOPD término que hace referencia a la Ley Orgánica de Protección de Datos.

Vaulting acción por la que se intercambia cintas de datos entre diferentes ubicaciones.

Retención periodo de tiempo establecido para salvaguardar la información.

5. Bibliografía

Informatica empresarial, Cuantic (2017). *31 de Marzo, día mundial del backup*. [online] Disponible en: http://cuantic.es/31-de-marzo-dia-mundial-del-backup/ [Accedido Febrero 2017].

Mularczyk, Sam (2017). *Get ready – World is March 31st!* [online] Disponible en: http://www.worldbackupday.com/es/ [Accedido Febrero 2017].

Amanda (2017). *Main Page*. [online] Disponible en: http://wiki.zmanda.com/index.php/Main_Page [Accedido Marzo 2017].

BackupPC (2017). *BackupPC*. [online] Disponible en: http://backuppc.sourceforge.net/ [Accedido Marzo 2017].

Bacula (2017). *Bacula*. [online] Disponible en: http://blog.bacula.org/ [Accedido Marzo 2017].

Bareos (2017). *Open Source Data Protection*. [online] Disponible en: https://www.bareos.org/en/ [Accedido Marzo 2017].

Commvault.com. (2017). *Enterprise Backup and Recovery | Backup & Recovery Solutions*. [online] Disponible en: https://www.commvault.com/itleaders [Accedido Marzo 2017].

Commvault (2017). *Commvault*. [online] Disponible en: https://www.commvault.com/ [Accedido Marzo 2017].

Commvault Systems. (19 de Enero del 2016). *Commvault Solution Guide*. [PDF] Disponible en: https://kapost-files-prod.s3.amazonaws.com/published/54466d4c096e133eff0006c4/virtualization-competitive-onesheet.pdf#page=1&zoom=auto,-206,798 [Accedido Marzo 2017].

Commvault Systems. (15 de Enero del 2015). *Commvault Licensing Information*. [PDF] Disponible en: http://www.arrowecs.es/web/producto/ofertas/marketing/2015/commvault/news_octubre/noticias/Commvault-Licensing-Information-Summaries.pdf [Accedido Marzo 2017].

IBM (2017). *IBM Spectrum Protect family*. [online] Disponible en: http://www-03.ibm.com/software/products/es/spectrum-protect-family [Accedido Marzo 2017].

VEEAM (2017). *Veeam*. [online] Disponible en: https://www.veeam.com/es/vm-backup-recovery-replication-software.html [Accedido Marzo 2017].

IBM (2017). *IBM Spectrum Protect family*. [online] Disponible en: http://www-03.ibm.com/software/products/es/spectrum-protect-family [Accedido Marzo 2017].

Microsoft, Azure (2017). *Calculadora de precios*. [online] Disponible en: https://azure.microsoft.com/es-es/pricing/calculator/?service=backup#backup-1 [Accedido Abril 2017].

Amazon, Web Service (2017). *Simple Monthly Calculator*. [online] Disponible en: http://calculator.s3.amazonaws.com/index.html [Accedido Abril 2017].

Google, Could Platform (2017). *Google Could Platform Pricing Calculator*. [online] Disponible en: https://cloud.google.com/products/calculator/#id=9348991e-9740-4176-911c-68e524760dc5 [Accedido Abril 2017].

Sarah, Mitroff (2017). *OneDrive, Dropbox, Google Drive and Box: Which cloud storage service is right for you?* [online] Disponible en: https://www.cnet.com/how-to/onedrive-dropbox-google-drive-and-box-which-cloud-storage-service-is-right-for-you/ [Accedido Abril 2017].

Muchmore, Michael and Duffy, Jill (2017). *The Best Cloud Storage and File-Sharing Services of 2017*. [online] Disponible en: http://www.pcmag.com/roundup/306323/the-best-cloud-storage-providers-and-file-syncing-services [Accedido Abril 2017].

Intronis [Intronis] (6 de Febrero del 2015). *The Intronis Reverse Incremental Backup Technique* [Archivo de video]. Disponible en: https://www.youtube.com/watch?v=9PqZkM6RerU [Accedido Mayo 2017].

Databarracks [Oscar Arean] (12 de Diciembre del 2013). *Incremental vs. Differential Backup* [Archivo de video]. Disponible en: https://www.youtube.com/watch?v=pmzeebcx-vk [Accedido Mayo 2017].

Agencia Española de Proteccion de Datos (13 de Diciembre del 1999). *Guia de la Seguridad de Datos*. [PDF] Disponible en: https://www.agpd.es/portalwebAGPD/canaldocumentacion/publicaciones/common/pdfs/guia_seguridad_datos_2008.pdf [Accedido Mayo 2017].

Wikipedia (25 de Febrero 2017). *Unidad de cinta*. [online] Disponible en: https://es.wikipedia.org/wiki/Unidad_de_cinta [Accedido Mayo 2017].

Wikipedia (30 de Junio 2017). *Cinta magnética de almacenamiento de datos*. [online] Disponible en: https://es.wikipedia.org/wiki/Cinta_magn%C3%A9tica_de_almacenamiento_de_datos

Ultrium LTO (1999). *What is LTO Technology?* [online]. Disponible en: http://www.lto.org/technology/what-is-lto-technology/ [Accedido Mayo 2017].

Ultrium LTO (1999). *FAQ* [online]. Disponible en: http://www.lto.org/about-the-lto-program/faq/ [Accedido Mayo 2017].

IBM Knowledge Center (1999). *WORM functionality for LTO tape drives and media*. [online]. Disponible en: https://www.ibm.com/support/knowledgecenter/en/STCMML8/com.ibm.storage.ts3500.doc/ipg_3584_meltoworm.html [Accedido Mayo 2017].

Hewlett Packard Enterprise (Diciembre 2015). *Encryption technology for HPE StoreEver LTO Ultrium Tape Drives*. [online]. Disponible en: https://www.hpe.com/h20195/v2/getpdf.aspx/4AA5-2801ENW.pdf?ver=Rev%201 [Accedido Mayo 2017].

Wikipedia. (Mayo 2017) *Hard disk drive*. [online]. Disponible en: https://en.wikipedia.org/wiki/Hard_disk_drive#Technology [Accedido Mayo 2017].

Pinola, Melanie (6 de Octubre del 2015). *Data Storage Technologies of the Future*. [online]. Disponible en: https://www.backblaze.com/blog/data-storage-technologies-of-the-future/ [Accedido Mayo 2017].

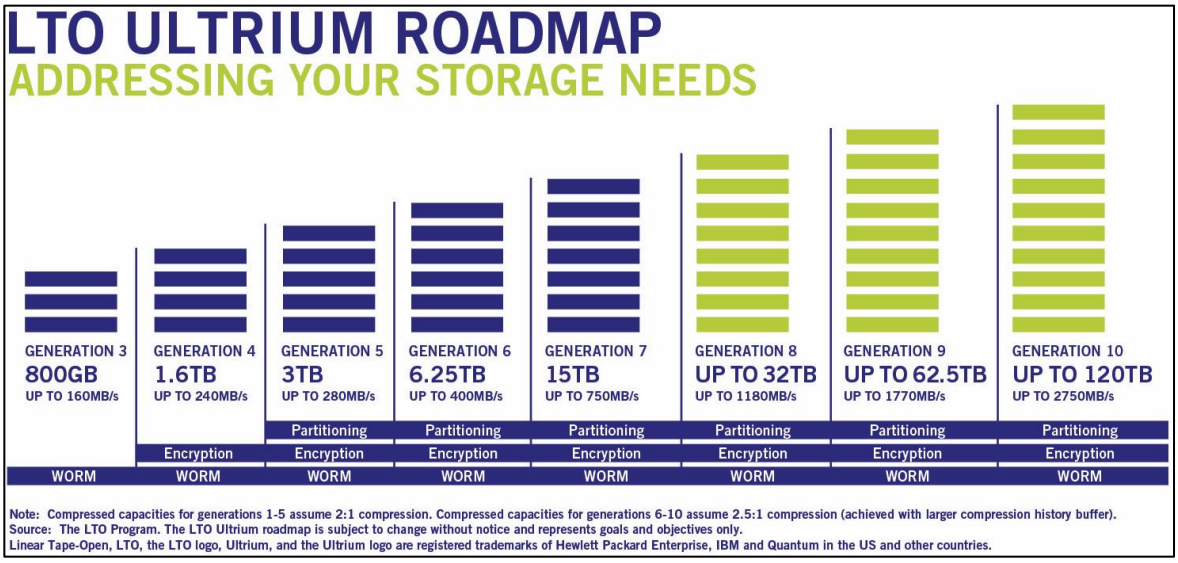
Solutions Squad (25 de Octubre 2013). *Different Types of backups explained*. [online]. Disponible en: https://www.youtube.com/watch?v=G1IdxM-r1fg [Accedido Mayo 2017].

CommVault (2016). Synthetic Full Backups. [online]. Disponible en: http://documentation.commvault.com/commvault/v10/article?p=features/backup/syn_full.htm

6. Anexos

6.1. Anexo I

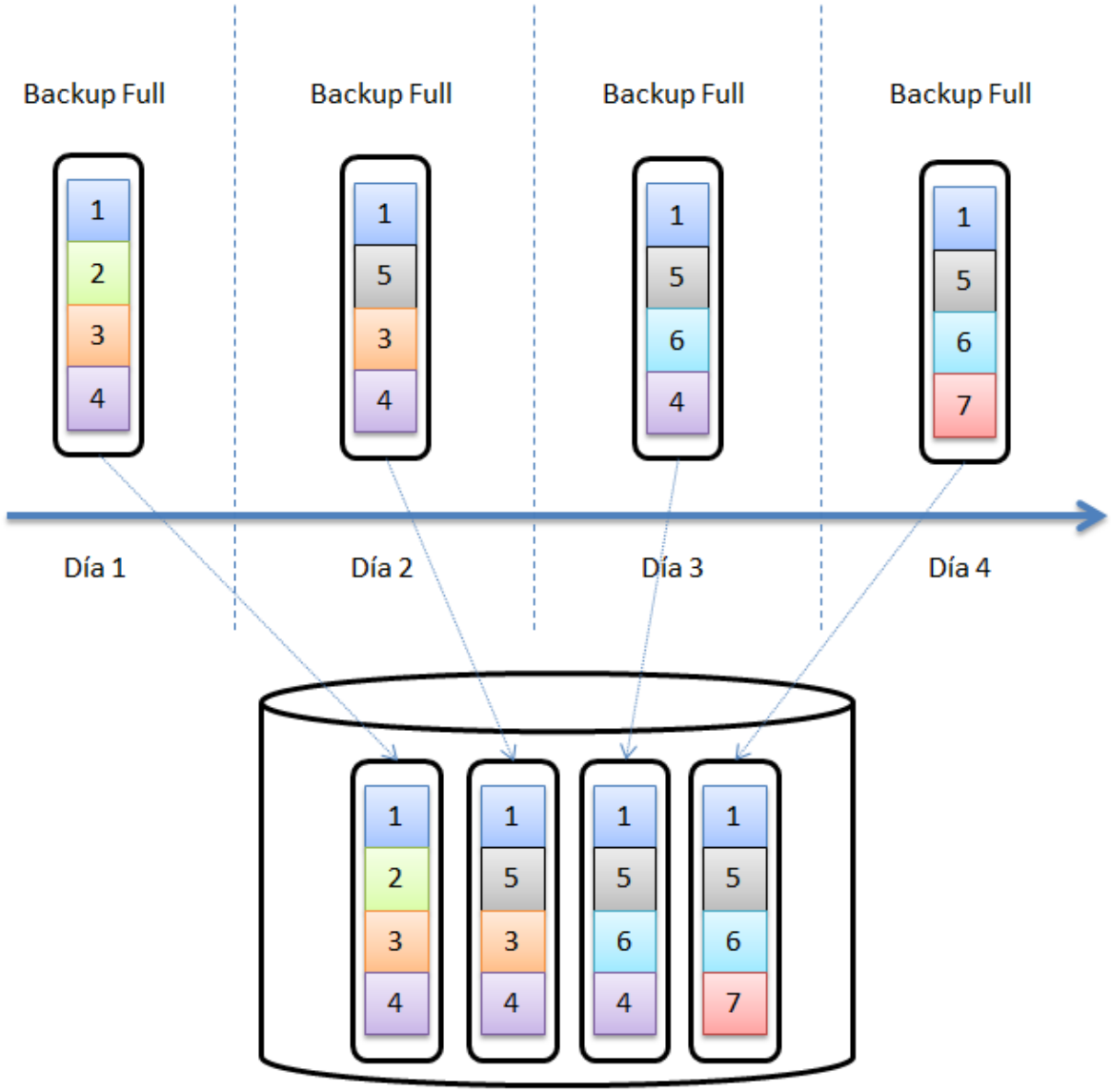
Ilustración 6.1-1 - LTO Ultrium Roadmap



Ultrium LTO (2017). *LTO Ultrium Roadmap*. [online] Disponible en: http://www.lto.org/wp-content/uploads/2014/06/LTO7_2C_10GenChart_0815_HPE_CMYK.png [Accedido Mayo del 2017].

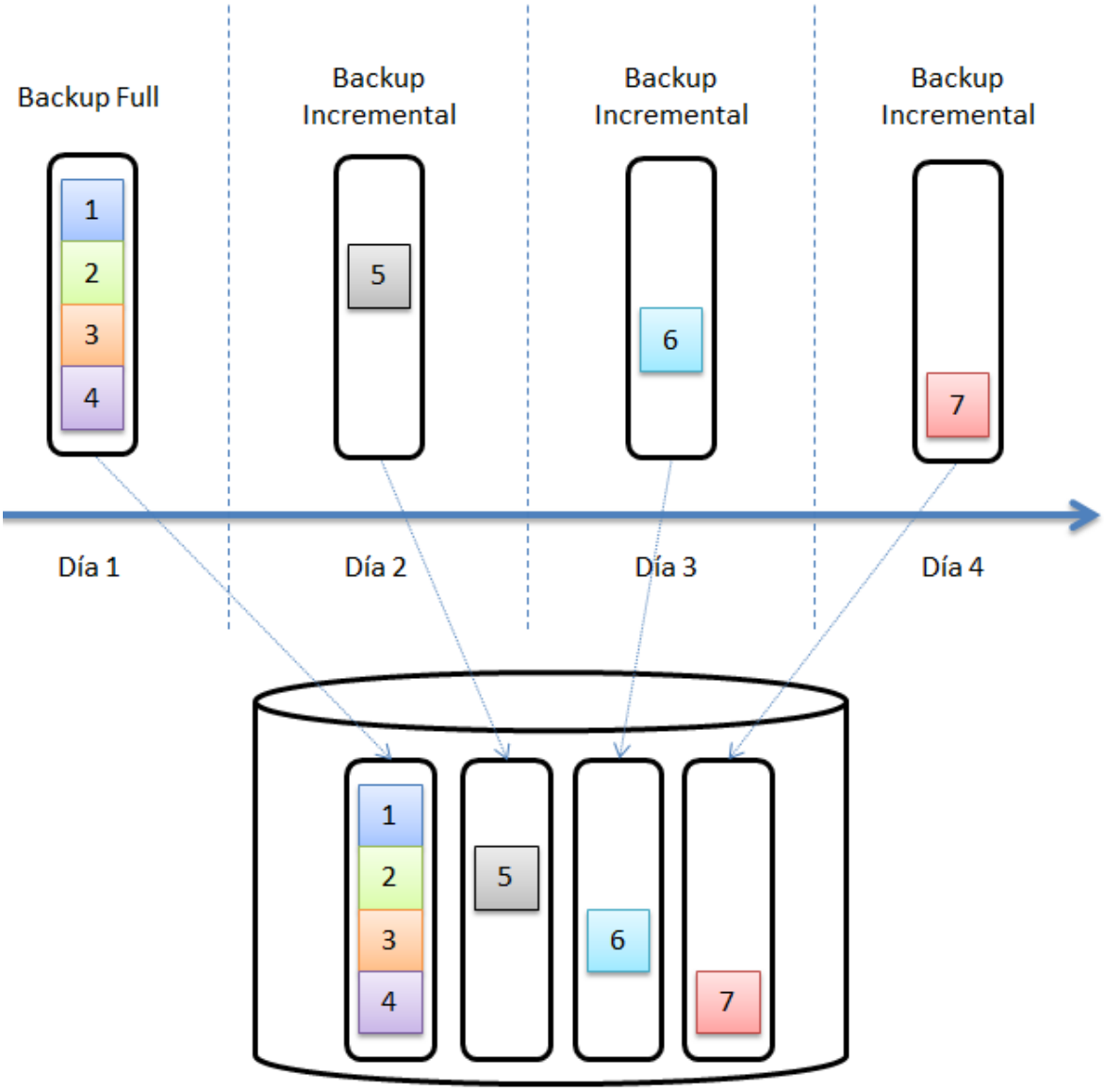
6.2. Anexo II

Ilustración 6.2-1 - Backup Full



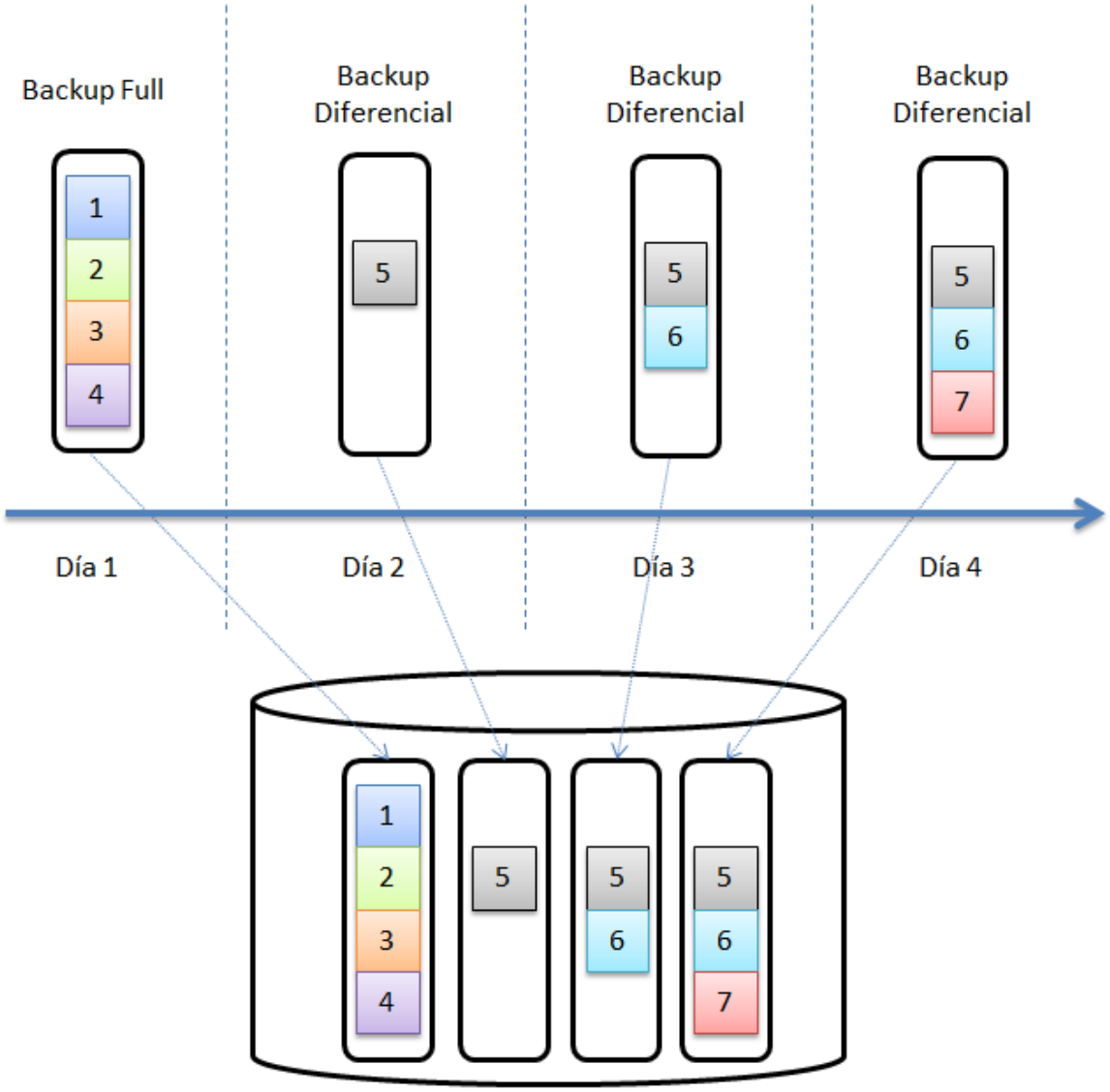
6.3. Anexo III

Ilustración 6.3-1 - Backup Incremental



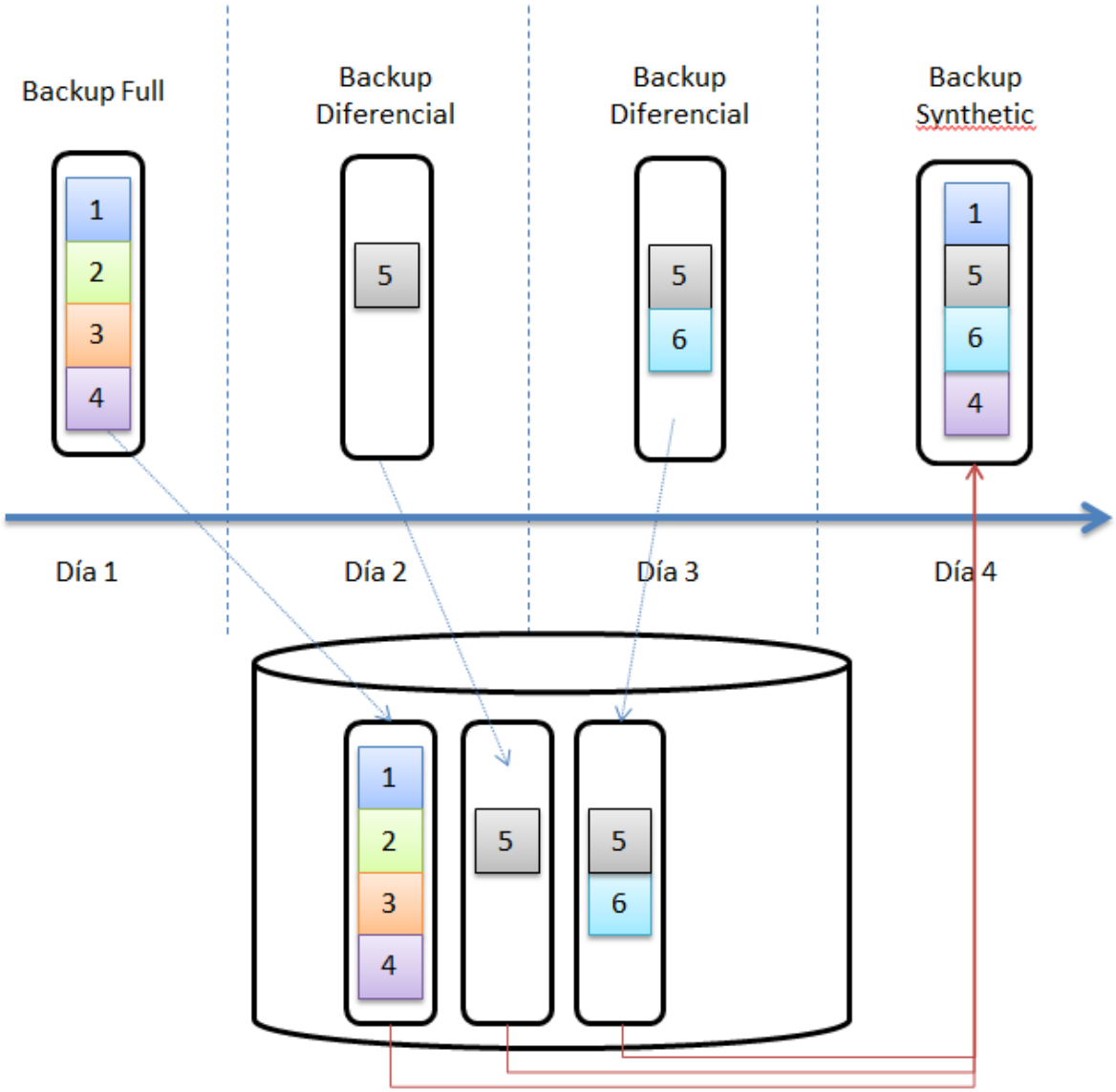
6.4. Anexo IV

Ilustración 6.4-1 - Backup Diferencial



6.5. Anexo V

Ilustración 6.5-1 - Backup Sintético



6.6. Anexo VI

Ilustración 6.6-1 - Backup Incremental Forever