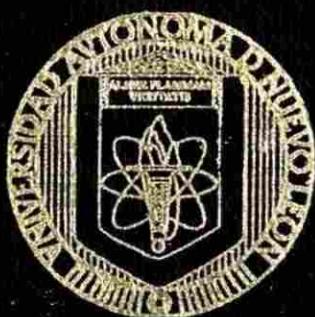


UNIVERSIDAD AUTONOMA DE NUEVO LEON

**FACULTAD DE CONTADURIA PUBLICA
Y ADMINISTRACION**

DIVISION DE ESTUDIOS DE POST-GRADO



**INTEGRACION Y ADMINISTRACION DE
DOCUMENTOS ELECTRONICOS:
HERRAMIENTA POTENCIAL EN LAS EMPRESAS**

POR

RICARDO JAVIER GUERRA DOMINGUEZ

**COMO REQUISITO PARCIAL PARA OBTENER EL
GRADO DE MAESTRIA EN ADMINISTRACION CON
ESPECIALIDAD EN SISTEMAS DE INFORMACION**

ENERO, 2002

2002

TM

Z7 1 6

FC PY

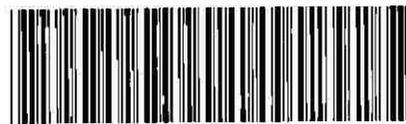
0002

0002

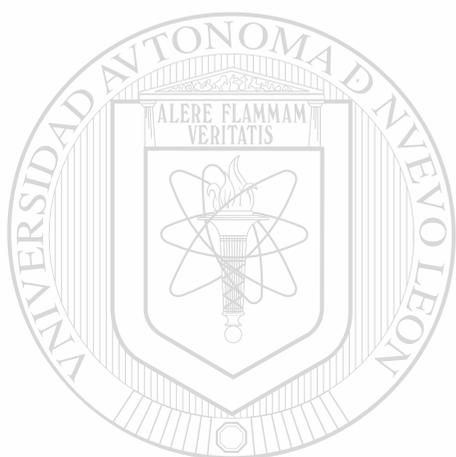
0002

0002

INTEGRAACION Y ADMINISTRACION DE
DOCUMENTOS ELECTRONICOS:
HERRAMIENTA POTENCIAL EN LAS EMPRESAS



1020146948



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

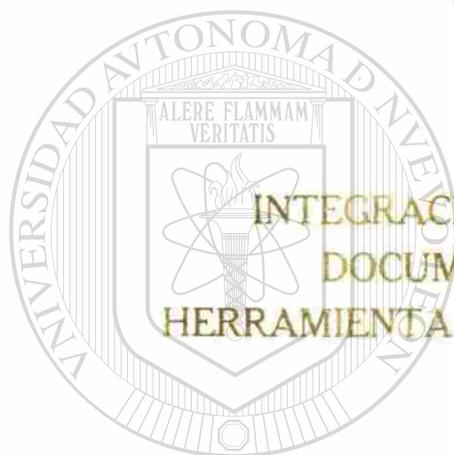


DIRECCIÓN GENERAL DE BIBLIOTECAS

UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE CONTADURIA PUBLICA
Y ADMINISTRACION

DIVISION DE ESTUDIOS DE POST-GRADO



INTEGRACION Y ADMINISTRACION DE
DOCUMENTOS ELECTRONICOS:
HERRAMIENTA POTENCIAL EN LAS EMPRESAS

POR

RICARDO JAVIER GUERRA DOMINGUEZ

UNIVERSIDAD AUTONOMA DE NUEVO LEÓN



DIRECCIÓN GENERAL DE BIBLIOTECAS

COMO REQUISITO PARCIAL PARA OBTENER EL
GRADO DE MAESTRIA EN ADMINISTRACION CON
ESPECIALIDAD EN SISTEMAS DE INFORMACION

ENERO, 2002

971216

TH

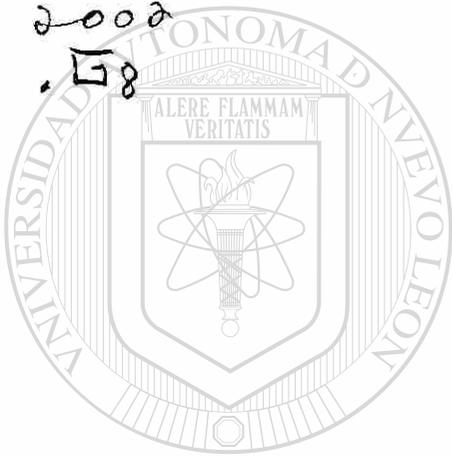
Z 7164

.C8

F0P4A

2002

.E8



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

®

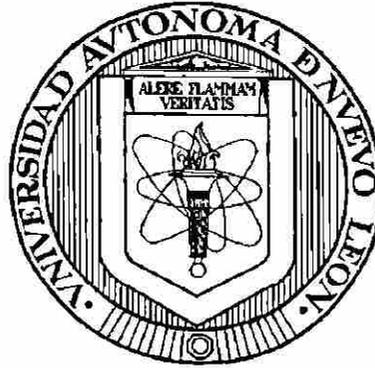
DIRECCIÓN GENERAL DE BIBLIOTECAS



FONDO
TESIS

UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE CONTADURIA PUBLICA Y ADMINISTRACION



**INTEGRACION Y ADMINISTRACION DE DOCUMENTOS
ELECTRONICOS :**

HERRAMIENTA POTENCIAL EN LAS EMPRESAS

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

Por
DIRECCIÓN GENERAL DE BIBLIOTECAS

RICARDO JAVIER GUERRA DOMINGUEZ

Como requisito parcial para obtener el Grado de
MAESTRIA EN ADMINISTRACION con Especialidad en
Sistemas de Información

Enero, 2002





UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

DIRECCIÓN GENERAL DE BIBLIOTECAS

®

AGRADECIMIENTOS

Al Lic. Enrique Jimenez Jimenez, Asesor de mi tesis, porque gracias a la gran experiencia con que cuenta en el asesoriamiento de tesis, pude terminar la investigación con los resultados deseados.

A la Lic. Patricia del Carmen Elguezabal Lopez e Ing. Jorge Alberto Méndez Dávila, por el apoyo recibido para el logro de esta meta.

Al C.P. José Magdiel Martínez Fernández, Secretario de Posgrado, por el apoyo y comprensión recibido en el proceso y ampliación en el proceso de presentación de la investigación.

A la escuela de Posgrado de FACPYA en general, desde el nivel inferior hasta el nivel de dirección, ya que ellos hacen posible el que cada vez mas personas logren superarse para alcanzar sus metas, a traves de la experiencia y conocimiento que te brinda esta misma.

DEDICATORIA

Gracias a Dios, ya que gracias a su bendición y protección, siempre he logrado las metas que he anhelado. Gracias, mil gracias Dios mio.

A mi abuelo Manuel, uno de los participes en la creación de la maravillosa familia que tengo, el cual admiro desde siempre por sus buenos consejos, su personalidad y su sabia forma de pensar.

A mi Madre, la cual con su temple y amor me ha guiado siempre por el buen camino, logrando así, día a día, alcanzar mis objetivos y metas con éxito.

A Belle, la principal motivación en mi vida, la mujer que me hizo descubrir que el amor llega en el momento exacto, con la persona que Dios nos tiene destinada.

A todas aquellas personas que de alguna manera me dieron muestras de apoyo durante el desarrollo de mi investigación.

INDICE

INTRODUCCION

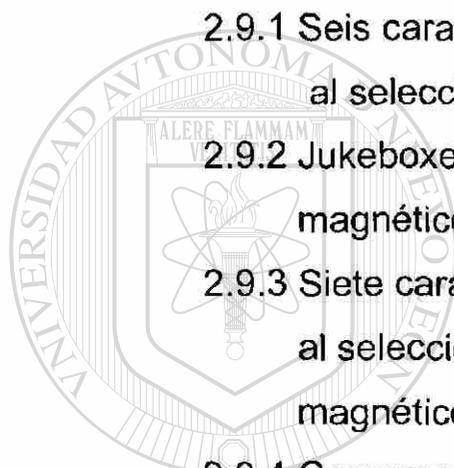
1. ESQUEMA GENERAL DE LA INVESTIGACION

1.1 Objetivo general de la investigación.....	1
1.2 Planteamiento del problema.....	2
1.3 Hipótesis.....	3

2. MARCO TEORICO

2.1 Antecedentes.....	4
2.2 Administración de archivo convencional.....	5
2.2.1 Documentos.....	5
2.2.2 Tipos de documentos.....	5
2.2.3 Control de archivo convencional.....	6
2.3 Control de espacio físico.....	6
2.3.1 Tipos de Administración.....	6
2.4 Sistemas para digitalizar documentos.....	7
2.4.1 Preparación de documentos físicos.....	7
2.4.2 Pasos para preparar documentos.....	9
2.5 Escaneo.....	10
2.5.1 Que se entiende por escaneo.....	11
2.5.2 Tipos de escanner.....	11
2.5.2.1 Tipos de configuración de escanner...	12
2.5.2.2 Interfaces de Software y Hardware más comunes para un escanner	13
2.6 El mejoramiento de una imagen, como puede ser aplicado, y que impacto tiene en el proceso de escaneo.....	14

2.7 Factores que provocan una pobre calidad de la imagen y como puede esto afectar a otros sistemas dentro de la organización.....	15
2.8 Estándares tecnológicos para el almacenamiento de documentos electrónicos.....	15
2.8.1 Identificación de los formatos de archivos mas comunes.....	16
2.9 Almacenamiento de documentos electrónicos.....	17
2.9.1 Seis características importantes a considerar al seleccionar un jukebox de CD/DVD.....	18
2.9.2 Jukeboxes para dispositivos ópticos magnéticos.....	19
2.9.3 Siete características importantes a considerar al seleccionar un jukebox de media óptico magnéticos.....	19
2.9.4 Comprando librerías de tape.....	20
<hr/>	
2.10 Dispositivos magnéticos.....	21
2.10.1 Discos Láser.....	22
2.10.2 Tape – Cartucho.....	22
2.10.3 Discos ópticos.....	23
2.10.4 Microfilm.....	23
2.11 WORM : Una opción para el acceso en línea.....	23
2.11.1 Necesidad de múltiples tipos de almacenamiento.....	24
2.11.2 El rol del almacenamiento óptico en las aplicaciones de digitalización de documentos.....	27



U A N L

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

DIRECCIÓN GENERAL DE BIBLIOTECAS



2.11.2.1 On-Line – En línea (On-Line).....	27
2.11.2.2 Near-Line.....	28
2.11.2.3 Far-Line.....	28
2.11.2.4 Off-Line.....	28
2.11.3 Alta capacidad Near-Line y un permanente almacenamiento de archivo.....	29
2.11.4 Cuando usar Near-Line	29
2.11.5 Borrables vs. WORM – La importancia de la permanencia.....	31

3. MARCO REFERENCIAL

3.1 Riesgos en el almacenamiento electrónico.....	33
3.1.1 Integridad de datos	33
3.1.2 Capacidad.....	35
3.1.3 Performance (Rendimiento)	37
3.1.3.1 Velocidad	38
3.1.3.2 Escribiendo	39
3.1.3.3 Leyendo	39
3.1.3.4 Durabilidad	41
3.1.3.5 Longevidad	43
3.1.3.6 Costos Actuales	45
3.1.4 Opciones de tecnología WORM	47
3.2 Comparando costos de compra vs. desarrollo de una aplicación completa de digitalización de documentos.	47
3.2.1 Fase de codificación	48
3.2.1.1 Lenguaje de programación	48
3.2.1.2 Conocimiento de la base de datos.....	49
3.2.1.3 Herramienta de desarrollo APIs.....	49
3.2.1.4 Revisión de errores.....	50

3.2.2 Fase de pruebas.....	51
3.2.3 Fase de documentación.....	52
3.2.4 Fase de mantenimiento.....	53
3.2.5 Fase de soporte.....	55
3.2.6 Cálculos finales.....	56

4. INVESTIGACION DE CAMPO

4.1 Caso Practico.....	58
4.2.1 FEDEX.....	58

5. ESTANDARES TECNOLOGICOS DE CLASE MUNDIAL

5.1 Formatos y tipos de compresión de una imagen.....	63
5.1.1 Comparación de formatos de imágenes convencionales.....	65

6. CONCLUSION

70

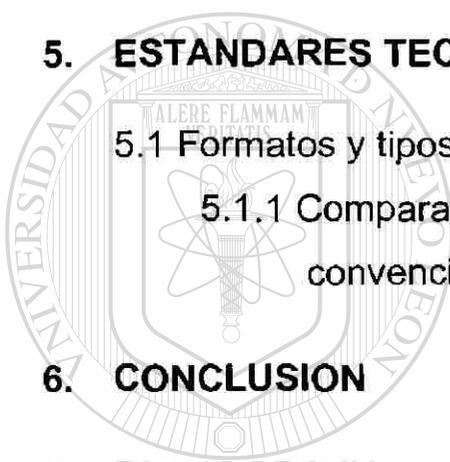
7. BIBLIOGRAFIA

75

8. APENDICE

76

A. Caso de la Embajada de Estados Unidos de Norteamérica en México: solución al manejo de documentación del proceso para obtención de la Visa.....	76
1 El problema	76
2 La solución	78
3 Un modelo de solución	79



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



B. Caso de la empresa Black&Decker aparatos eléctricos almacena mas que taladros y cortadoras de césped 80

 1 La compañía..... 80

 2 El reto..... 80

 3 La solución..... 81

 4 Retorno de la inversión..... 82

 5 Resumen de beneficios..... 82



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



DIRECCIÓN GENERAL DE BIBLIOTECAS

Introducción

En una empresa, se crea un archivo físico cuando se requiere guardar documentos generados de esta misma por distintos motivos, los cuáles pueden ser para efectos fiscales, administrativos, históricos, etc. . Al paso del tiempo se comprueba que, el tener un archivo físico equivale a costo, ésto por requerir de espacio físico, recurso humano y tiempo para administrar este mismo. Así también se tiene la probabilidad de riesgo de pérdida de documentos por factores ajenos a la empresa, como pueden ser los factores naturales (humedad, tiempo), factores accidentales (incendio, robo) entre otros.

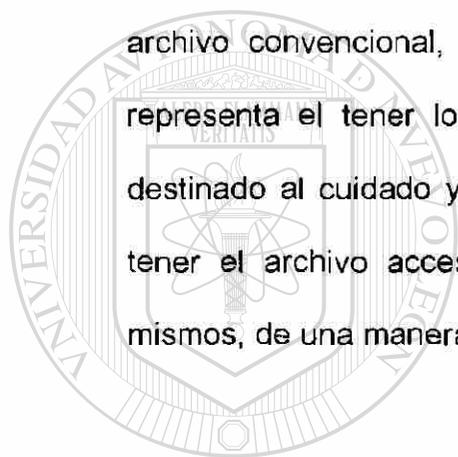
El archivo electrónico es una solución para la optimización de costo y tiempo, así como para la reducción de riesgo de pérdida de información. Ésto se realiza por medio del escaneo de documentos para convertir el

documento físico en imágenes. Pero, ¿tienes conocimiento de que tipo de escanner debes de utilizar?, ¿Que, dependiendo del tipo de papel se determina qué configuración se debe de definir tanto para el escanner como para la aplicación?, ¿Que hay varios tipos de almacenamiento para las imágenes de los que comúnmente conoces?, ¿De los pasos que se deben de seguir en el proceso de digitalización de documentos?, ¿De los riesgos que puedes tener durante el proceso?, ¿De los factores clave para determinar si se compra la aplicación o se desarrolla en la misma empresa?, etc. . Todo esto se describe en los siguientes capítulos que comprende esta investigación.

1. ESQUEMA GENERAL DE LA INVESTIGACIÓN.

1.1 Objetivo general de la investigación.

Se realizará una investigación para encontrar la manera de reducir al máximo los costos administrativos, el espacio físico y el tiempo requerido para acceder la información, que comprende el manejo y mantenimiento del archivo convencional, esto es: eliminar el costo de espacio físico que representa el tener los archiveros, reducir el costo del recurso humano destinado al cuidado y administración de dicho archivo, así como también tener el archivo accesible al personal autorizado para el uso de estos mismos, de una manera rápida y efectiva.



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

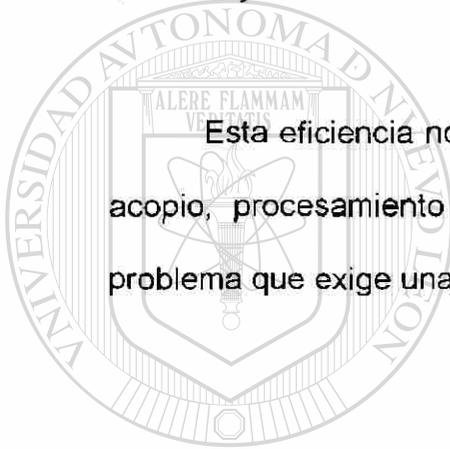
®

DIRECCIÓN GENERAL DE BIBLIOTECAS

1.2 Planteamiento del problema.

Las empresas del tercer milenio con presencia en todo el mundo; generan, procesan y utilizan volúmenes de información cada vez mayores; y por otro lado, el mismo entorno exige permanentemente un acceso más eficiente y efectivo a la información como soporte de la toma de decisiones.

Esta eficiencia no se puede lograr con los esquemas tradicionales de acopio, procesamiento y acceso a la información, generándose así un problema que exige una solución rápida y certera.



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



DIRECCIÓN GENERAL DE BIBLIOTECAS

1.3 Hipótesis.

Es necesario crear un nuevo concepto de informática en las empresas que manejan grandes volúmenes de documentos e información, que se traduzca en un modelo de sistema de almacenamiento, proceso y acceso de información necesaria para las empresas que deseen estar dentro de los estándares de las empresas de clase mundial.

Este nuevo concepto consiste en lograr generar un sistema de procesamiento de documentos, de tal manera que cualquier persona de una empresa, pueda acceder la información que esta misma requiera, de una manera rápida y sencilla, con el fin de soportar una toma de decisión profesional. El sistema abarcará recursos y estándares internacionales, los

cuales son avalados por organizaciones oficiales reconocidas a nivel mundial.

DIRECCIÓN GENERAL DE BIBLIOTECAS

A partir de lo anterior, pretenderemos demostrar dicho sistema con el presente trabajo.

2. MARCO TEÓRICO.

2.1 Antecedentes.

Todas las organizaciones requieren concentrar información, procesarla y acceder oportunamente y de manera confiable para tomar decisiones.

No obstante, en muchas de estas organizaciones no logran reunir la información necesaria en el momento oportuno, aún y cuando dicha información se encuentra en su mismo archivo.

Aunado a lo anterior, se necesita tener mucho espacio físico para almacenar el archivo, así como también es necesario tener personal encargado del cuidado y administración de este mismo.

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN[®]
DIRECCIÓN GENERAL DE BIBLIOTECAS

2.2 Administración de archivo convencional.

2.2.1 Documentos.

El documento es todo el papel el cuál contiene conceptos predeterminados por una empresa, con el fin de ser escritos por la persona interesada. Dicho documento por lo general es almacenado en un archivo convencional.

2.2.2 Tipos de Documentos.

Actualmente existen por lo general archivos convencionales, de los cuales se derivan varios tipos de documentos que se deben de guardar por un tiempo determinado, dependiendo tanto de políticas de la empresa, así

como del gobierno:

Documento de control interno .-

- Tarjetas de control de inventario.

- Facturas de la compañía.

- Archivo departamental (recursos humanos, etc.).

Documentos externo.-

- Hacienda(*).

- Auditoria externa.

(*) Hay que recordar que la legislación de la materia exige a los contribuyentes conservar la información hasta por cinco años.

3.2.3 Control de archivo convencional.

La administración del archivo convencional varía según la cantidad de documentos, tamaños, tipos o nivel de consulta, y dependiendo de lo anterior la empresa determina si se almacena en bodegas externas, en el mismo edificio, o incluso, en alguna oficina específica.

El control de acceso a los documentos se hace sobre la base de sencillas formas de entrada y salida, con lo cual se busca evitar el extravío de cualquier documento.

2.3 Control de Espacio Físico.

2.3.1 Tipos de Administración.

La cantidad de personas que administran el archivo convencional depende del tamaño de la empresa, esto es, si se trata de una empresa pequeña, solo se necesita de una persona para llevar el control, pero hablando de una empresa de talla nacional o internacional, las cuales manejan volúmenes extensos de documentos, se pueden necesitar de entre 2 y 6 personas. Dichas personas se encargan de llevar el control de entradas y salidas de documentos, así como verificar que las cajas, archiveros, etc., estén en condiciones óptimas para evitar el maltrato de la documentación contenida en estos mismos, por lo cual, en caso de ser

necesario, dicho personal se encarga de la sustitución de las cajas o archivero dañados por otro en buen estado.

2.4 Sistemas para digitalizar documentos.

De acuerdo a la referencia del punto anterior, en la administración de archivos convencionales, es necesario crear otro archivo de tipo electrónico, y no obstante, aunque en la actualidad existe la tecnología y equipo necesario, falta la conceptualización y concreción de un nuevo tipo de archivo. Esto se va a conformar con:

- Nuevos depósitos de información.
- Nuevos esquemas de proceso de información.
- Nuevos sistemas de disposición de información.

2.4.1 Preparación de documentos físicos.

Este proceso no es tan simple como se ve, pues existen muchos factores acerca de las condiciones físicas del documento, así como el determinar el tipo de organización que se hará a éste mismo.

El documento o papelería tiene diferentes formas de presentación como son los tamaños, pesos, densidades, acabados y colores, y éstos a su vez pueden contener grapas, clips, cinta adhesiva, con los cuales puede estar aun más deteriorado el documento. Estos factores afectan el tiempo total de la preparación de los documentos, y esto a su vez envuelven los factores de decisión acerca de que scanners se usarán para digitalizar los documentos. Esto también afecta el tiempo de digitalización, la velocidad de

conversión de documento físico a documento electrónico, así como también los requerimientos de control de calidad, indexación y almacenamiento.

†

Las características de un documento son :

a) Dimensiones.

Las mediciones tanto horizontal como vertical ayudan a determinar el tipo de manejo del mecanismo que puede ser usado.

b) Peso.

A mayor peso del papel, mayor es el grueso de éste mismo. El papel grueso tiende menos a ser maltratado, y por ende es menos transparente. La transparencia o el papel muy delgado puede ser un problema para las condiciones de luz intensa que se genera dentro del escáner, esto si hay algo escrito detrás del papel a escanear.

c) Terminado del papel.

El alto grado de terminación y brillo del papel, al ser ambas reflejadas en la luz del escáner, se ven tan diferentes como la comparación de una revista y un papel periódico. La terminación del papel afecta directamente la escanabilidad de este mismo. Esto también impacta en que tan bien puede pasar la hoja por el alimentador automático de papel, esto en el escáner.

d) Color.

El papel de color, y la naturalidad de cómo el color está plasmado en el papel, impacta como el papel se ve para el ojo humano, más aún, como es visto por el escáner.

A partir de estas cuatro características se determina que tipo de escanner se deberá utilizar, así como el proceso de preparación de los documentos mismos.

2.4.2 Pasos para preparar documentos.

Este punto es uno de los más importantes en la etapa de la digitalización del documento, puesto que es la base para poder saber que información se encuentra y en qué documento.

a) Armado de documentos .-

Es el ordenamiento y acomodo de los documentos según su importancia, tamaño y estado físico en el que se encuentre. Así también se realiza la acción de quitar cualquier elemento físico que pueda dañar el escanner, o bien el mismo documento. Dichos elementos pueden ser grapas, clips, broches, pegamento, etc. .

b) Identificación del documento .-

Es la identificación oficial que tendrá cada documento a escanear. Dicha identificación se hace a partir de un nombre único del documento, seguido de una clave única de éste mismo, el cual identificará el tipo de

documento y a que persona, oficina, empresa, corporativo, etc., le corresponde dicho documento, para utilizarlo en el sistema a integrar.

2.5 Escaneo.

Para empezar a hablar del escaneo quiero citar una anécdota de una persona, y esto fué que cuando tenía 10 años, un amigo le dejó un bumerán de madera. Él había oído tremendas historias sobre los peligros de esta arma legendaria, pero, como era un chico aventurero, quería descubrir por el mismo todas las historias que había oído acerca del bumerán. Sin apenas hacer esfuerzo, lanzó el bumerán tan alto y tan lejos que dejó de verlo. Al voltearse con otro amigo, emocionado con lo que pasó, le platicó como le había hecho y siguieron jugando a otro juego, en eso, lo único que recuerda es que él estaba tirado en el suelo con una herida ensangrentada en la nuca, mirando confundido al cielo y preguntándose que era lo que lo había golpeado.

Los usuarios novatos, algunos con cierta experiencia y, a veces, incluso los profesionales de los dispositivos del escaneo sufren a menudo una situación similar. Los scanners son herramientas muy potentes para reproducir imágenes del mundo real en formato digital, y sus tecnologías de conectar y usar (plug and play) hacen que el proceso de capturar datos visuales sean muy sencillo. Ponga una hoja con la cara de la hoja hacia abajo, baje la tapa, pulse sobre el botón escanear del programa, y listo. Después coloque las imágenes resultantes en un archivo de composición de página y veamos lo que sale.

Hasta aquí todo bien. Pero cuando se quiere escanear documentos más grandes que el escáner existente en la compañía, o que cuando quiera imprimir la imagen salga con un aspecto granulado y con colores terrosos, o cuando los documentos legales o dibujos salen con una nitidez tan baja que resulta imposible leerlos, uno se siente tentado a pensar que el arte de escanear es una cuestión de acertar o fallar.

Las buenas noticias son que existen estrategias que permiten asegurar una alta calidad a las imágenes escaneadas, sea cual sea el producto final.

2.5.1 Que se entiende por escaneo

Cuando utilizamos el término escaneado (digitalización), lo que realmente queremos decir es adquisición de imágenes. También llamada entrada digital, la adquisición de imágenes es el proceso de capturar información visual en formato digital, apto para que las computadoras puedan utilizarla.

2.5.2 Tipos de Escáner

Actualmente existen, para escanear documentos, los siguientes tipos de escáner :

- 1- De plataforma.
- 2- De alimentador de hojas.
- 3- Para diapositivas.
- 4- De tambor.
- 5- Cámaras digitales.
- 6- Sistemas de procesado de foto CD.

Pero para efectos de nuestro estudio se mencionarán aspectos de los primeros 4 tipos de escáner, aunque en la actualidad, hay escáners (scanners) con dos o más tipos integrados, como ejemplo de ésto, es un escáner que tiene plataforma y a su vez alimentador de hojas.

2.5.2.1 Tipos de configuración de escáner.

Cuando se van a escanear documentos, se debe de tener conocimiento acerca de los dispositivos de hardware para escanear que deben de tener, así también se debe de tener conocimiento de los componentes físicos del escanner así como su funcionalidad.

También se deberá tener conocimiento del software que puede interactuar con el hardware, con el fin de ver la capacidad asociada que se tiene entre ellos, y ver como puede ser usada durante el escaneo de documentos .

En orden para configurar un sistema de escaneo, se necesita seleccionar la arquitectura de los componentes para el escaneo de imágenes que debe ser usada, y como pueden ser utilizadas en conjunto con eficiencia.

Los escanners utilizan una variedad de interfaces de hardware y software, así como dispositivos especializados de software que son creados específicamente para cierto tipo de equipo.

2.5.2.2 Interfaces de Software y Hardware más comunes para un escanner.

Las interfaces del escanner caen en dos categorías:

- Interfaces de hardware.
- Interfaces de software.

a) Interfaces de hardware.

Las interfaces de hardware más comunes para los escanners son SCSI (Small Computer System Interfase) y las interfaces de videos. La interface de video provee a la estación de escaneo una salida directa de CCD (dispositivo de conexión de carga), el cual se encuentra dentro del escanner, y este a su vez debe estar conectado directamente a una tarjeta de procesamiento de imágenes, la cual convierte la señal análoga en señal digital para dar finalmente por resultado la imagen del documento. Un escanner con conexión SCSI procesa la señal internamente en la tarjeta.

Así también, existen escanner de rendimiento bajo, los cuales cuentan con interfaces paralelo o serial, y éstas pueden ser conectadas a una interface paralela/serial en cualquier computadora personal.

b) Interfaces de software.

Mientras que las interfaces de hardware proveen la habilidad para la conexión física de los dispositivos del escanner, las interfaces de software proveen una conexión lógica. Las interfaces lógicas más comunes que existen son ISIS y TWAIN, las cuales son el interpretador de la comunicación entre el escáner y la computadora.

2.6 El mejoramiento de una imagen, como puede ser aplicado, y que impacto tiene en el proceso de escaneo.

El mejoramiento de la imagen es una técnica que mejora la calidad de la imagen en orden para proveer el menor tamaño de la imagen sin eliminar o dañar la información importante o más solicitada en un documento escaneado. Los mejoramientos de imágenes típicos son :

a) Detección del Borde de la hoja (Edge detection).

Es la habilidad de remover la información no requerida en las orillas de la imagen.

b) Corrección de inclinación (Deskewing).

Corrige electrónicamente la corrección de inclinación de la hoja, ocasionada durante el escaneo de esta misma.

c) Smoothing.

Remueve las líneas causadas por las arrugas del papel, así también remueve puntos que se encuentran en las orillas afiladas.

d) Rotación.

Es la habilidad de rotar una imagen, la cual pudo ser escaneada al revés, es decir, pasando primero la parte de abajo del documento.

Dichas capacidades pueden ser de hardware, el cual puede procesarse en un escanner con interface SCSI. Para los escanners con tarjeta de video, dichas capacidades son procesadas en la tarjeta de procesamiento de imágenes que reside en la computadora a la que esta conectada el escanner. Por último, cabe mencionar que dichas capacidades

de mejoramiento de imagen pueden ser realizadas mediante puro software sin la intervención de hardware, pero el tiempo de proceso depende de la velocidad y capacidad de la computadora personal, ya que se procesa en la memoria de la computadora.

2.7 Factores que provocan una pobre calidad de la imagen y como puede esto afectar a otros sistemas dentro de la organización.

Los factores que pueden afectar la calidad de una imagen son la pobre preparación del documento, constantes cambios o cambios inadvertidos en la resolución del escanner, la mala selección del hardware o software a utilizar. Lo anterior puede resultar en la necesidad de re-escanear los documentos, en una ineficiente búsqueda de información, en un ineficiente almacenamiento de imágenes, también puede originar un excesivo tráfico en la red.

2.8 Estándares tecnológicos para el almacenamiento de documentos electrónicos.

Una imagen es comprimida y almacenada con un formato de archivo para que pueda ser administrada por una aplicación. Esto es importante saber, porque de esto depende el rango del tamaño del archivo, de acuerdo al formato de archivo y algoritmo seleccionado.

Diferentes tamaños de documentos, escaneados con diferentes resoluciones y comprimidos a diferentes radios de compresión, crean un rango ancho de tamaños de archivos que impactan el rendimiento de un

sistema. El tamaño de la imagen tiene un impacto significativo en varios subsistemas, así como en la capacidad de almacenamiento requerido, y la eficiencia y velocidad de una red computacional. El proceso de arquitectura del proceso de escaneo debe entender estas implicaciones al momento de configurar los sistemas y redes relacionadas con este mismo.

2.8.1 Identificación de los formatos de archivo más comunes.

El algoritmo de compresión mas común es el CCITT Grupo-III y el Grupo-IV, así también el LZW y el JPEG.

a) Compresión CCITT Grupo-III y Grupo-IV.

Este tipo de compresión es optimizado para compresiones bitonales, es decir, comprime en blanco y negro. El Grupo-III es un estándar de compresión análoga y el Grupo-IV es un estándar de compresión digital.

b) LZW.

Este tipo de compresión es optimo para la compresión de caracteres de datos.

c) JPEG (Join Photographic Experts Group).

Este estándar de compresión es optimo para documentos complejos y fotografías.

Los ratios de compresión varían de imagen a imagen, y pueden ser comprimidos en el rango de 10:1 o menos para la compresión de Grupo-III, y 20:1 o menos para la compresión de Grupo-IV.

2.9 Almacenamiento de documentos electrónicos.

Toda documentación electrónica se debe de guardar en dispositivos magnéticos, determinando la empresa el requerido de acuerdo a sus necesidades y capacidades tecnológicas.

Afortunadamente, la tecnología esta respondiendo a la insaciable demanda de velocidad y capacidad, así como librerías y jukeboxes para almacenamiento masivo.

Las aplicaciones de documentación intensiva tienen una gran confianza en las soluciones de almacenamiento near-line por tener un rápido acceso al gran volumen de información archivada. Mientras que los discos magnéticos ópticos alguna vez dominaron por un tiempo, ahora existen mas opciones económicas disponibles, que incluyen el DVD y las librerías de tape para manejo de memoria en el cartucho, las cuales ofrecen un acceso rápido a los datos.

DIRECCIÓN GENERAL DE BIBLIOTECAS

No obstante, antes de tomar la importante decisión acerca de la compra de algún dispositivo de hardware específico, se deben de considerar, para cada tipo de hardware, las siguientes características, las cuales no se deben de omitir al realizar dicha operación :

2.9.1 Seis características importantes a considerar al seleccionar un jukebox de CD/DVD.

- a) Formato .- Si estas considerando utilizar DVD, es importante saber si tu demanda de almacenamiento requerido es para escribir solo una vez, o reescribir varias veces en el mismo DVD.
- b) Tiempo para el cambio de CD/DVD .- La velocidad del robot y la distancia que viaja el CD/DVD dentro del jukebox determina el tiempo promedio en cambiar la media. Un doble brazo para la selección de la media puede también incrementar el tiempo de cambio de este mismo.
- c) Capacidad vs. velocidad.- ¿Tu prioridad es maximizar el almacenamiento o minimizar el tiempo de consulta?, Esta pregunta te ayudara a determinar la cantidad de dispositivos para disco. Si tu tienes muchos usuarios, teniendo mas dispositivos ayudara para asegurar un rápido acceso a la información.
- d) Tamaño del cache o memoria interna.- Un jukebox con un gran cache volverá dramáticamente mas lento el tiempo de respuesta para los accesos frecuentes a documentos.
- e) Soporte para CD y DVD .- ¿El jukebox soporta dispositivos DVD-RAM, DVD-ROM o DVD-R?, La mayoría de los jukeboxes ofrecen diferentes configuraciones que te permiten aumentar el numero de dispositivos o módulos de discos de acuerdo a las necesidades de cambio de almacenamiento que se requiera.
- f) Rápida alternancia.- Algunos jukeboxes permiten a un disco sencillo leerlo a través de un batch de lectura de los módulos de discos mientras la unidad esta en línea.

2.9.2 Jukeboxes para dispositivos ópticos magnéticos.

Cuando el almacenamiento de información es sujeto a estándares regulatorios, esto por parte de comisiones de seguridad e intercambio o alguna otra agencia reguladora, no existe algo igual a la confiabilidad y aseguramiento del almacenamiento óptico magnético. La mayoría de las medias ópticas magnéticas tienen un promedio de vida de 30 años o más, la velocidad de respuesta es inigualable, y la opción WORM (escribir una vez leer muchas) te da la seguridad de almacenamiento de información al no poder ser borrada esta misma.

2.9.3 Siete características importantes a considerar al seleccionar un jukebox de media óptico magnéticos.

- a) Capacidad.- ¿Puedes agregar más discos cuando las necesidades de almacenamiento de incrementen?. Algunos jukebox Ópticos magnéticos son modulares y pueden ser actualizados.
- b) Costo.- No solo vea el costo del hardware. El software, la instalación y el servicio de contrato pueden elevar el costo.
- c) Radio de distancia entre el cartucho y el manejador de este mismo.- Busca un radio de distancia corto o la opción de agregar más manejadores en base al incremento de discos.
- d) Robot.- ¿La unidad utiliza buscadores sencillos o dobles?, Al usar un buscador doble, reduce el tiempo de intercambio de discos, esto por traer dos discos en lugar de uno solo.

- e) ¿Tiempo de acceso a los datos.- Cual es el tiempo requerido que se tiene en mente para ver los datos?. El tiempo de lectura te proporciona una mejor comparación entre dos o más jukeboxes.
- f) Confiabilidad.- Identificar que opciones de servicio y soporte ofrecen las distintas compañías productoras de jukeboxes.
- g) Compatibilidad.- ¿La unidad o unidades de jukebox compatibles?.
¿Podrás actualizar fácilmente el manejador de cierta cantidad de Gbytes a uno de mayor capacidad el siguiente año?.

2.9.4 Comprando librerías de tape.

Las preguntas y dificultades que se detallan a continuación deberán tomarse en cuenta al momento de comprar lectores y librerías para tapes.

- a) Conexiones de red.- ¿Qué interface utiliza la librería?, ¿Es compatible con tu red?.
- b) Manejadores.- ¿Cuántos manejadores ofrece la librería?, ¿Podrá ser posible hacer actualizaciones a nuevos formatos de tape?
- c) Escalabilidad.- ¿Puedes combinar múltiples librerías para incrementar la capacidad cuando las necesidades de almacenamiento lo requiera?, ¿Se puede incrementar el numero de manejadores o gabinetes?, ¿Cuánto costaría una actualización?.

- d) Robot.- Una vez que hayas seleccionado varios modelos para escoger uno, compara el tiempo promedio de movimiento del robot y el tiempo promedio de lectura del manejador.
- e) Lectores de código de barras removibles o fijos. ¿El lector de código de barras es estándar u opcional?, Los códigos de barras en los tapes disminuyen el tiempo necesitado para encontrar la información. Los lectores de código de barras movibles disminuyen el uso del tape porque escanean estos mismos mientras los cartuchos siguen en su lugar.
- f) Servicio/garantía.- ¿Qué garantía, servicio y soporte son incluidos?, ¿Si el fabricante de la librería está usando manejadores de otras compañías, quienes le pueden dar soporte a estos?, ¿Que tiempo de garantía ofrece la compañía?
- g) Futuros tipos de lectura .- ¿Existen controles de seguridad para la lectura o manejo de tapes?

2.10 Dispositivos magnéticos.

Son los dispositivos en donde se puede almacenar cualquier tipo de información electrónica, con el fin de tener segura la información de una empresa.

Existe una variedad de dispositivos magnéticos y recientemente no magnéticos, los cuales se seleccionan sobre la base de las necesidades de la empresa:

2.10.1 Discos Láser.

a) CD-ROM.

Este disco esta hecho premasterizado, es decir, no se puede rescribir o borrar información del disco.

b) CD-R.

Es un CD-ROM en blanco en el cual se puede grabar información en sesiones, es decir, si se graba cierta información y se cierra la sesión, la única manera de seguir escribiendo en este dispositivo es creando una nueva sesión.

c) CD-RW.

Este dispositivo magnético te permite escribir y borrar en la misma área del disco las veces que se necesite.

2.10.2 Tape - Cartucho.

Son dispositivos elaborados de cintas magnéticas, las cuales pueden ser de distintos tamaños, tanto en longitud de la cinta como en dimensiones del cartucho. Actualmente existen cartuchos de distintos tamaños, los cuales son : Cartucho y Mini-Cartucho.

2.10.3 Discos Ópticos.

a) WORM.

Las iniciales WORM describen el significado en inglés de Write Only Read Many. Este tipo de disco óptico permite a las imágenes a escribirse una sola vez y poder leerse muchas, por lo cual el almacenamiento se hace de manera incremental.

2.10.4 Microfilm.

Son fichas plásticas en las cuales se graba la información una sola vez. El acceso a la información se vuelve lenta y tiende a ser un archivo físico pero en menor escala.

2.11 WORM : Una opción para el acceso en línea.

Almacenamiento óptico puede ser una tecnología poco familiar a muchas personas. Aun así, es la opción del ideal por usar en numerosas aplicaciones que envuelven números inmensos de documentos y grandes cantidades de datos que demandan alta capacidad, seguridad y almacenamiento accesible a un precio económico. Estas aplicaciones tienen que guardar y administrar millones o billones de documentos. (Como se escucha, billones de documentos con "B").

Ciertos tipos de problemas no son resueltos por una típica aplicación de negocios, ni aquellos que requieren el tipo común de soluciones del almacenamiento. Aquí es donde WORM (write once/read many) que es un tipo de almacenamiento óptico, es una solución tecnológica muy eficaz.

Normalmente referido a por las siglas WORM, este tipo de solución de almacenamiento óptico ha sido una opción atractiva para las aplicaciones documentación de imágenes de gran escala.

El tamaño más popular para el WORM es la tecnología del platter de almacenamiento óptico de 12 pulgadas de diámetro. La tecnología del platter de 12-pulgadas ofrece muchas ventajas y beneficios que son importantes para muchas aplicaciones de misión crítica en el que negocios y gubernamentales cuentan.

2.11.1 Necesidad de múltiples tipos de almacenamiento.

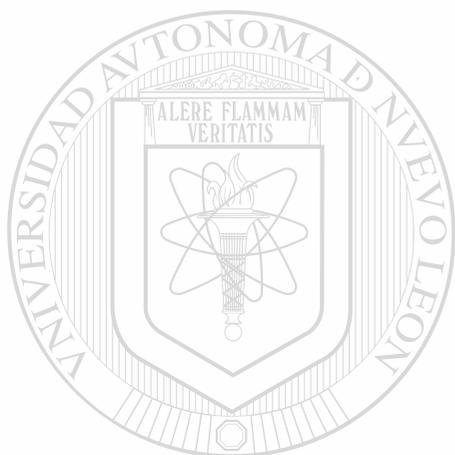
Los requerimientos de almacenamiento en el medio ambiente computacional de la mayor parte de los corporativos van bien mas allá de cualquier tipo de almacenamiento o tecnología. De cualquier manera, la gran cantidad de tecnología disponible hace que la selección ideal del proceso de almacenamiento de información se convierta en una tarea difícil de hacer.

Este proceso puede ser simplificado entendiendo el concepto de la herencia de almacenamiento. Este modelo se agrupa esencialmente en grupos de almacenamiento de acuerdo a su capacidad y performance.

Generalmente, existe una división entre capacidad y performance global. Los dispositivos que te ofrecen un rápido acceso a la información

comúnmente se caracterizan por su baja capacidad y alto costo, mientras que los dispositivos que son relativamente de bajo costo, gran capacidad de almacenamiento, ofrecen un almacenamiento lento y un bajo performance.

A continuación se presenta una gráfica, en la cual se detalla el comportamiento de los factores costo, capacidad y velocidad, esto de acuerdo a las necesidades y requerimientos tecnológicos en una aplicación.



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

®

DIRECCIÓN GENERAL DE BIBLIOTECAS

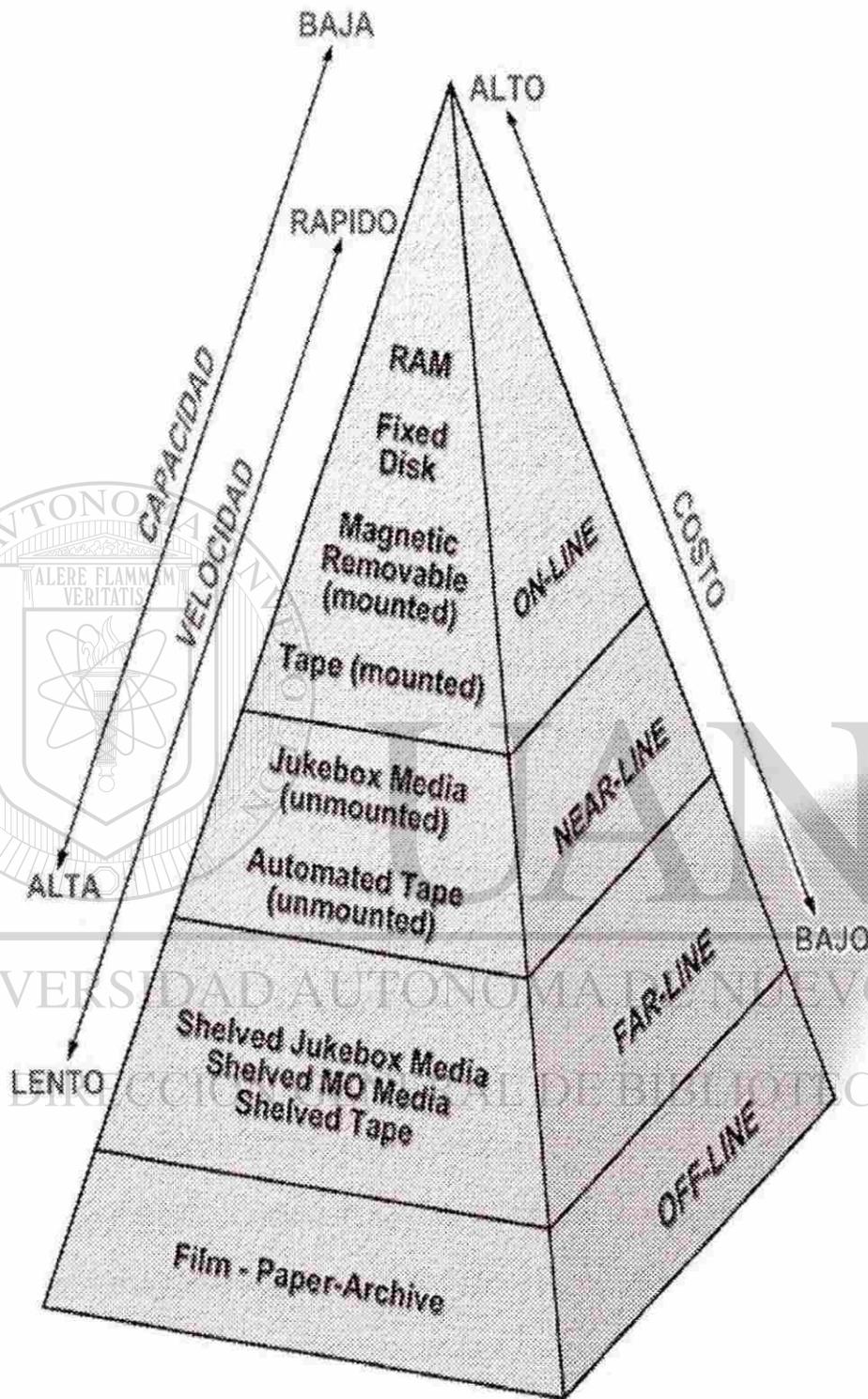


Tabla 1 .- Factores Costo, Capacidad y Velocidad en dispositivos de almacenamiento.

El objetivo es encontrar el mejor precio, performance, y capacidad para las necesidades de una corporación, usualmente a través de la combinación de tecnologías de almacenamiento.

2.11.2 El rol del almacenamiento óptico en las aplicaciones de digitalización de documentos.

Cuando las imágenes de los documentos necesitan estar almacenadas por periodos largos de tiempo, estos pueden estar almacenados en cierto medio de almacenamiento durante los ciclos de proceso, otros documentos en medios para acceso intermitente a través de periodos de tiempo, y otros en medios de retención de periodos largos y archivo. El almacenamiento óptico puede jugar un rol importante a través de la mayoría de los ciclos de vida del documento.

Los actuales tipos de almacenamiento son usados de acuerdo a la ubicación lógica en la cual el documento digitalizado debe residir a través de las diversas fases de actividad.

DIRECCIÓN GENERAL DE BIBLIOTECAS

2.11.2.1 On-Line - En línea (On-Line).

Se refiere como el almacenamiento el cual da disponibilidad de acceso inmediata mediante un sistema computacional, y este se encuentra en un dispositivo de lectura/escritura. Los documentos digitalizados son almacenados en línea cuando estos necesitan ser accedidos instantáneamente. Este tipo de requerimiento es normalmente solucionado en el almacenamiento en el disco duro magnético, que en la actualidad es típicamente puesto en una configuración de alta capacidad RAID

(Redundant Array of Independent Disks). Los requerimientos del almacenamiento en línea pueden ser también solucionados por dispositivos de accesos variables, incluyendo discos ópticos montados en manejadores de almacenamiento óptico.

2.11.2.2 Near-line .

El almacenamiento Semi-en-línea (Near line) puede ser accedido en segundos sin la intervención humana en dispositivos como la librería robótica de almacenamiento o mejor conocido como "jukeboxes". Los documentos digitalizados son almacenados semi-en-línea cuando el nivel de tiempo en traer un documento digitalizado es aceptable para el tiempo de respuesta de una aplicación. Los documentos digitalizados en un disco no montado dentro de un jukebox son considerados semi-en-línea (near-line).

2.11.2.3 Far-Line.

El almacenamiento en-línea-lejana es cuando la media requiere intervención humana para que dicha media sea montada, justo como un plater de disco óptico tiene que ser almacenada en un estante.

2.11.2.4 Off-Line.

La Media o documento digitalizado que no son controlados por el sistema exportador son considerados consultas fuera-de-línea (Off-line). En ciertos momentos, el documento digitalizado puede ser movido fuera de su lugar habitual, o migrado a un sistema diferente, o bien a un archivo permanente.

La mayor parte de las implementaciones de almacenamiento jerárquico requerirán una combinación de tipos de media. El configurar un almacenamiento jerárquico es complejo para identificar la capacidad requerida y configuraciones específicas para cada tecnología, y la configuración del sistema entero con los distintos tipos de media combinados, así como los parámetros para mover los archivos de un tipo de media a otro.

2.11.3 Alta capacidad Near-Line y un permanente almacenamiento de archivo.

Mientras el almacenamiento óptico puede ser usado en línea como se menciono anteriormente, los medios ópticos, especialmente la media que solo se puede escribir una vez (write-once), es típicamente usado para Near-line y archivo permanente. Los discos magnéticos es normalmente la media que se selecciona para el almacenamiento en línea (on-line) en la mayoría de las aplicaciones. Esto es debido a las características del disco magnético, incluyendo el alto rendimiento, capacidad de rescribir, tipo de configuración de arreglos, etc..

2.11.4 Cuando se debe de usar Near-line.

El propósito principal del almacenamiento near-line es satisfacer las demandas de respuesta que no necesariamente debe estar o residir en línea (on-line), pero debe de ser accesible en un periodo de tiempo requerido; sin embargo, la respuesta a la pregunta "¿Cuanta capacidad se

debe determinar para el almacenamiento near-line?", Es simple, pues es el satisfacer los requerimientos de acceso a los datos activos con un aceptable tiempo de respuesta, el cual normalmente se da por segundos. Este tiempo varia de aplicación en aplicación. Cada aplicación tiene sus propias características y requerimientos de performance; sin embargo, cada aplicación deberá tener diferentes tipos de respuesta a las siguientes preguntas :

- a) ¿Que periodo de tiempo se debe de mantener los documentos en un nivel de actividad de recuperación, sin abarcar el periodo de tiempo que justifique el costo del almacenamiento en línea?
- b) ¿Hay una necesidad de recuperar ciertos tipos de documentos mientras alguien espera en la línea telefónica?.
- c) ¿Podría una operación beneficiar estando habilitada para recuperar documentos mientras alguien este esperando en el teléfono, o en persona?
- d) ¿Cuales son los tamaños y volúmenes de documentos que requieres de acceso sin intervención humana?
- e) ¿Cuál es el nivel de tiempo de recuperación aceptable para ciertos tipos de documentos?
- f) ¿Cuál es el crecimiento esperado en cuanto a volumen de documentos en los siguientes par de años?
- g) ¿Que tan importante es para la organización el alcanzar o tener el total de datos habilitados en un sistema automatizado?

h) ¿Cuál es el impacto global operacional al no tener ciertos documentos automáticamente accesibles?

2.11.5 Borrables vs. WORM - "La Importancia de la Permanencia".

La pregunta acerca del uso de dispositivos WORM(write once-read many) contra el uso de dispositivos rescribirles puede tener muchas respuestas.

Para algunos, es fácil contestar la respuesta - "¿Puede una operación crear el riesgo de tener ciertos documentos sobre-escritibles, borrados o alterados?", y , "¿Son estos requerimientos regulatorios a considerar para usar dispositivos write-once para ciertos documentos?".

Algunos documentos caen en la categoría de registros permanentes, en le verdadero significado de la palabra registro, como evidencia, prueba, confirmación, testimonio o testigo. En estos casos, la permanencia de ciertos registros o documentos deben estar seguros. El inalterable dispositivo WORM nos da ese nivel de permanencia a escribir la información o documento digitalizado, al menos hasta que el dispositivo pueda ser leído, dependiendo de la duración de vida del dispositivo magnético.

Otros documentos pueden ser clasificados como "estáticos" con requerimientos de retención específico de tiempo. En algunos casos, los documentos digitalizados deben solo permanecer intactos hasta alcanzar el fin del periodo de retención, y después de este punto, estos pueden ser borrados o destruidos.

Para algunos documentos, el periodo de retención puede estar entre el rango de un par de años hasta varias décadas, y es así como se asigna el uso de una permanente solución de almacenamiento WORM; de cualquier manera, cuando el periodo de retención es solo por un periodo corto (1 o 2 años por ejemplo), este puede ser sensible a simular un almacenamiento tecnológico WORM con dispositivos reusables o re-escribibles.

Sin embargo, la simulación del almacenamiento WORM trae riesgos, esto porque tanto los dispositivos como los datos son protegidos solo por software a través de switches y métodos físicos de protección contra escritura, como el aditamento que tiene un disquete de 3 1/2" para proteger contra escritura. En casos donde los documentos son actualizados continuamente en un periodo de tiempo y finalmente se hacen estáticos en un lapso de tiempo, una solución apropiada es la combinación de tecnología de almacenamiento borrable y el almacenamiento óptico WORM.

DIRECCIÓN GENERAL DE BIBLIOTECAS



3. MARCO REFERENCIAL.

3.1 Riesgos en el almacenamiento electrónico.

3.1.1 Integridad de Datos.

No existe pregunta acerca de cual es el recurso de mayor valor en cualquier sistema computacional, - este es la información. Así también, la información es el más difícil recurso a reemplazar. La mayoría de las compañías y organizaciones gubernamentales, grandes y pequeñas, podrían hasta parar sus operaciones por varios días si estas tuvieran una pérdida de información mayor. La importancia de la integridad de la información se incrementa día con día sobre la base del crecimiento económico global y viene a ser más dependiente sobre la información digitalizada en la compañía.

Los tipos de dispositivos de almacenamientos tienen sus variaciones de vulnerabilidad. RAM (Random Access Mode) por ejemplo, es totalmente volátil. Tan pronto como el equipo computacional donde se encuentra el RAM es apagado, toda la información en el RAM desaparece. Pero aun algunos dispositivos "no-volátiles" tienen su grado de vulnerabilidad.

Mientras los llamados discos (magnéticos) duros no son vulnerables para cualquier tipo de propósito, estos son altamente propensos a que se dañen de algún sector. Los pasos para recuperar información a partir algún daño del disco o alguna falla en el mecanismo, el disco tiene que ser

reemplazado y se deberá recuperar la información a partir de un dispositivo de respaldo, en el cual en la mayoría de los casos es un tape magnético. Un nivel de protección contra el paro de operaciones por una falla en el disco duro se puede dar implementando un sistema RAID, pero aun así, esto no protege al disco duro de fallas mayores como son el fuego, sabotaje o desastres naturales.

Los dispositivos de almacenamiento óptico no son vulnerables y este no es propenso a fallas de discos duros causados por el mecanismo de las cabezas lectoras de lectura / escritura. Esto es especialmente verdadero en un inalterable dispositivo WORM porque solo se escribe una vez la información o imágenes, y estas no pueden rescribirles o borrarles. A diferencia de los tapes y discos magnéticos, con almacenamiento óptico no hay contacto físico entre las cabezas lectoras y el dispositivo. La lectura y escritura son realizadas mediante láser grabando capa por capa dentro del plástico dura con una capa protectora, esto para proteger mejor contra daños físicos o daños por el medio ambiente que los rodea.

A diferencia de los dispositivos de discos magnéticos que guardan permanentemente la fuente de discos, las fuentes de almacenamiento óptico son removibles y portables. Esto hace que las estrategias de recuperación contra desastres sean fáciles de implementar. Las "copias espejo" de las fuentes ópticas pueden crearse, y guardar una copia en una parte cercana y otro respaldo enviarlo a alguna parte lejana y segura. Para el caso de un desastre mayor, las fuentes de respaldo simplemente llegan a ser las

fuentes primarias, y la información puede ponerse disponible sin un procedimiento de recuperación largo.

3.1.2 Capacidad.

La capacidad de almacenamiento, en cualquier tipo de dispositivo, se duplica por cada 1-3 años debido a los avances tecnológicos. Al mismo tiempo, la necesidad de una alta capacidad de almacenamiento continua incrementándose a un promedio asombroso sobre la base de las nuevas aplicaciones que son implementadas.

Referente a la digitalización de documentos, las imágenes bitmap son considerablemente más grandes que sus contrapartes electrónicas. El tamaño de una simple página tamaño carta es de aproximadamente 2Mbytes. La misma página, cuando esta es escaneada a 300 Dpi(puntos por pulgada), y comprimida con el tipo de compresión CCTT Grupo IV/2D es aproximadamente de 20:1 por lo tanto debe de estar entre el rango de 50Kbytes. Por lo tanto, hay que tener claro la alta capacidad de almacenamiento cuando queremos retener las imágenes por un periodo de tiempo.

Solamente calcula cuanta capacidad de almacenamiento se requiere para el procesamiento de 40,000 imágenes por semana dentro de una organización, en la cual cada semana debe almacenarse por 7 años y debe estar en línea-cercana (near-line) para un rápido acceso. Primero, el

porcentaje del tamaño de la imagen debe ser imaginado, por lo cual vamos a decir por ejemplo, que el promedio es de 60Kbytes.

$40,000 \times 60\text{Kbytes} = 2,400,000 \text{ Kbytes}$ o 2.4Gbytes por semana

$2.4\text{Gbytes} \times 52 \text{ Semanas} = 124.8 \text{ Gbytes}$ por año

$124.8 \text{ Gbytes} \times 7 \text{ años} = 873.6 \text{ Gbytes}$ como almacenamiento total.

Si estos volúmenes permanecen constantes, un jukebox de almacenamiento óptico con una capacidad de 900 Gbytes podría satisfacer el requerimiento de la aplicación.

La siguiente pregunta es: ¿Cuántas piezas de dispositivos de almacenamiento óptico debe tener el jukebox?. Si el requerimiento total de almacenamiento es de 873.3Gbytes, nosotros podemos calcular los dispositivos requeridos dividiendo el total de almacenamiento entre la capacidad de almacenamiento de cada tipo de dispositivo. A partir de esto, se dan las siguientes capacidades disponibles para los diversos tipos de dispositivos de almacenamiento óptico del tipo Write-once.

Tipo de Media	Capacidad	Cantidad requerida para almacenar 873.6 GB
CR-R	650 MB (un solo lado)	1344
DVD-R	3.95 GB (lado sencillo)	222
5.25"	5.2 GB (doble lado)	168
Optical		
12"	30 GB (doble lado)	30
Optical		

Tabla 2.- Cantidad de media requeridos para almacenar 873.6GB.

El número de dispositivos que se calcularon para los 873.6 Gbytes puede ser ligeramente mayor para evitar una sobresaturación de espacio.

Así también, la cantidad de dispositivos puede ser el doble por el motivo de la creación de una copia de respaldo para efectos de una recuperación de la información en caso de desastre. Finalmente, en caso de que se necesite un requerimiento de conversión cuando se llama a una imagen, esto puede incrementar la capacidad global requerida en nuestro ejemplo.

Recordando lo anterior, los 40,000 documentos digitalizados por semana se suman a los más de 14.5 millones de páginas sobre el periodo de 7 años. Muchos tipos de documentos requieren que se retengan o guarden por mucho tiempo, y algunos se tienen que guardar permanentemente. Estos requerimientos son el porqué muchas organizaciones terminan con warehouse (almacén) del tamaño de un estadio de fútbol solamente para almacenar documentos físicos que posiblemente nunca vayan a ser requeridos.

3.1.3 Performance (Rendimiento).

Mientras algunas aplicaciones pueden tolerar muchos días para responder a peticiones realizadas por usuarios, las organizaciones de todo tipo se sienten presionadas en el tiempo de entrega o respuesta, sino inmediato, rápido. Este es en los casos imperativos en donde se busca mejor el servicio

al cliente, particularmente en cualquier aplicación donde el tiempo de respuesta de la transacción impacta directamente al usuario.

Incluso las agencias de gobierno están presionadas en los procesos de respuesta de todos los tipos de transacciones, desde los ingresos de impuestos, acceso a registros públicos, hasta permisos y licencias.

Una gran capacidad de almacenamiento de dispositivos removibles hace factible el mantener grandes volúmenes de documentos en forma digital, así se proveerá una búsqueda y respuesta rápida. De esto hay tres áreas claves donde el performance necesita ser considerado: Efectividad en la velocidad de lectura y escritura, durabilidad en los dispositivos y medios ópticos, y longevidad de los medios ópticos.

3.1.3.1 Velocidad.

Mientras el promedio de transferencias es especialmente importante en escrituras y lecturas en grandes volúmenes, la efectividad en la velocidad para escribir y leer incluye mas que solo el promedio de transferencia del dispositivo. Existen una serie de pasos necesarios para una respuesta completa de peticiones cuando se está leyendo y escribiendo. Estos pasos dependen en el mayor de los casos en los dispositivos individuales y los controladores de estos. Dichos pasos incluyen

3.1.3.2 Escribiendo.

- Preparación del archivo
- Toma de demanda de escritura
- Identificación de la locación
- Transferencia del archivo
- Hacer disponible el medio magnético
- Mejoramiento de la escritura
- Verificación
- Actualización de archivo de tablas o bases de datos
- Posible depuración de un archivo de una locación previa
- Preparación del siguiente archivo

3.1.3.3 Leyendo.

- Recepción de la petición
- Identificación de la locación
- Poner disponible el medio magnético
- Mejoramiento de la lectura
- Transferencia del archivo
- Estar listo para la próxima petición

La tabla que a continuación se muestra contiene los promedios de tiempo de acceso y el rango en cuanto a Cantidad de bytes-tiempo durante una transferencia de información para los distintos tipos de dispositivos de Write-Once.

Tipo de Media	Tiempo de Acceso	Rango de transferencia
12"	25-100ms	6.0MB/s(lectura) y 2.5MB/s(escritura)
5.25" WORM	25-90ms	Entre 4.5MB/s en lectura y 2MB/s en Escritura
CD-R	150-450ms	150-1200KB/s en escritura (1-8x) y arriba de 4.8MB/s en lectura(32x)
DVD-R	Entre 100ms	Entre 1.4MB/s en escritura Y entre 1.5 a 8 MB/s en lectura
Tape	Entre 8 y 55 segundos en promedio	De 1.5 a 10 MB/s en lectura y escritura

Tabla 3 .- Tabla de rendimiento de dispositivos ópticos.

El rendimiento para cada paso puede variar sobre la base de las diferentes tecnologías de almacenamiento. Los dispositivos removibles pueden tener un rendimiento menor en la lectura y escritura en comparación a los discos duros magnéticos los cuales oscilan en un rango por abajo de entre 10ms a 20ms de tiempo de acceso, y en transferencia de datos puede alcanzar un promedio de 40MB/s.

De cualquier manera, el rendimiento efectivo de las dos tecnologías recién mencionadas no tienen mucha diferencia, esto una vez que la media de almacenamiento óptico es montada en el manejador. Las velocidades y promedio de transferencia se continúan mejorando, esto en la media de

almacenamiento óptico, y en si, actualmente son mejores que algunos discos magnéticos que se utilizaban no mucho tiempo atrás.

Con el almacenamiento near-line, la mayoría de los retardos provienen de la actividad mecánica del brazo del robot del jukebox. La media necesita ser constantemente sacada, montada, desmontada, lo que equivale a tiempo. La más alta capacidad de la media, es el menor numero de platters (fuentes) requeridos, esto porque minimiza el numero de cambio de platters.

Como se observa en el ejemplo anterior, la considerable alta capacidad de la media de almacenamiento óptico de 12" reduce significativamente el numero de platters necesarios, esto en comparación con otros tipos comunes de almacenamiento que son de medias removibles.

3.1.3.4 Durabilidad.

Sin embargo, lo que se dicta para determinar la durabilidad y longevidad de las medias es la combinación de materiales usados para la construcción de la misma media, y en un análisis final, se define que los factores principales son el manejo, mantenimiento y la acción de almacenamiento.

La media removible debe soportar la exposición al medio ambiente, así como el manejo tanto humano como mecánico. Entre mejor protegidas estén las medias, mayor tiempo de vida tendrán estas. En una sistema de librería manejado por un robot, por ejemplo, la media de almacenamiento de

12" tiene un rango de cientos de miles de accesos de carga y descarga y millones de lectura. Esto no se puede decir para todos los tipos de media, como el tape por ejemplo, que solo puede manejar varios miles de accesos de carga y descarga, así como miles de lectura-escritura.

Las medias CD y DVD pueden ser durables, tanto tiempo como no sean expuestas con frecuencia, así como el tomar en cuenta el cuidado del manejo humano. A diferencia de las medias de almacenamiento óptico de 5.25" y 12", el CD y DVD no tienen caja fija de protección, sin embargo, esto hace al CD y al DVD más susceptible de un daño físico, el cual puede comprometer la información. La caja protectora con un diseño para dar durabilidad conserva al disco óptico a ser tocado por el humano o dispositivos mecánicos.

Sin embargo, una mayor protección para el CD y el DVD puede ser optimizado manteniéndolos dentro de un jukebox para casos en que la información de estos dispositivos necesite estar en near-line, o bien deben de estar en estuches para cuando la información no se requiera en línea.

En resumen, seamos sinceros, en cuanto al manejo de los CDs y DVDs, nadie procura siempre manejar dichas medias agarrándolas de las orillas, o del centro, como lo recomienda la OSTA (Asociación de almacenamiento óptico tecnológico) así como los fabricantes de dichas medias.

3.1.3.5 Longevidad.

Nadie quisiera encontrarse ellos mismos en la situación de tratar de leer cierta información crítica, para descubrir que esta no esta. Afortunadamente, con los avances de las medias actuales nosotros podemos estar seguros que la información podrá estar a salvo mínimo por varios años en una media de almacenamiento segura. De hecho, la media de almacenamiento y la información contenida en este mismo pueden estar mas a salvo que el equipo y programas requeridos para el acceso, lectura e interpretación de la información almacenada.

Aunque pruebas sofisticadas pueden mejorar y proveer un indicador del tiempo de vida estimado de la media, ninguno puede dar con certeza, que el tiempo de vida que tendrá la media es la que resultó a partir de las pruebas.

Aun así, la mayoría de los fabricantes ofrecen garantías en la vida de la media en dos formas: Sin tener información almacenada y teniendo información almacenada, obviamente la más importante es el tiempo de vida de la media con información almacenada.

Porque nadie puede realmente asegurar cuan larga es la vida de la media, no solo debemos tener respaldos y plan de recuperación contra desastres, si no también es recomendable hacer una migración de la información a una nueva y actual tecnología. Estos métodos de protección pueden ser simplificados por unas pocas piezas de media y de alta

capacidad, que para ambos casos, son características de la media de almacenamiento óptico de 12".

La tabla que a continuación se muestra, resume el promedio de números obtenidos de varios fabricantes de medias, esto para cada tipo de media, así como pruebas de duración realizadas por el Laboratorio Nacional de Media, en San Paul Minesota. El ancho rango del periodo de vida se debe a la posible variación de calidad de cada fabricante, la combinación de materiales, así como los factores de operación y medioambiente de almacenamiento, principalmente la temperatura y la humedad relativa.

Cuando la temperatura y la humedad relativa son bajas, las expectativas de vida de la media es significativamente extendida, como puede verse en la siguiente tabla.

Tipos de Media	Tiempo de vida
12"	De 30 a 200 años
5.25" WORM	De 30 a 100 años
CD-R	De 10 a 100 años
DVD-R	De 10 a 100 años
Tape o Cartucho	De entre 5 y 30 años

Tabla 4 .- Expectativa de vida de la media óptica

La durabilidad también depende de los factores del medio ambiente en donde la media es almacenada. Los fabricantes comúnmente indican los

requerimientos en sus hojas de especificaciones, para medio ambientes operantes y no operantes; como ejemplo tenemos cuando se especifica la temperatura, ésta debe oscilar entre 0 y 20 grados Celsius. Adicional a esto, la media no debe ser expuesta a la luz solar directamente por periodos largos de tiempo, así también debe ser almacenada con niveles de humedad específicas, y debe conservarse limpia usando herramientas y materiales apropiados para la limpieza de esta misma.

En realidad, la media con una expectativa de vida mayor a 20 años sobrevivirán, la mayoría de ellas, a los sistemas de almacenamiento y lectura. Los cambios tecnológicos y el avance de esto son inevitables, por lo tanto la migración de la información a un una nueva media y un nuevo sistema de almacenamiento cada par de años debe ser necesario, esto para preservar los registros que requerimos almacenar.

3.1.3.6 Costos Actuales.

Es fácil hacer una estimación de costos por megabyte entre un tipo de media y otro, pero no es nada fácil el comparar con exactitud el verdadero costo que le pertenece por megabyte a un nivel de subsistema de almacenamiento. Para tener más exactitud en los resultados, la comparación de costos necesitan ser hechos a través de un periodo de tiempo de muchos años, como por ejemplo, de 3 a 7 años en un sistema real propietario. El mejoramiento del análisis de costos del subsistema envuelve muchas áreas de costos diversas.

- a) El costo de entrada por todo el subsistema de almacenamiento, incluyendo todos sus componentes como manejadores, jukeboxes, periféricos dedicados (servidores), programas para la administración de almacenamiento, y las medias para un periodo de tiempo, esto incluyendo las medias de respaldo.
- b) El costo de mantenimiento, incluyendo los contratos de servicio al equipo físico, el mantenimiento a los programas, el soporte a los sistemas operativos, actualizaciones de programas, y pequeñas actualizaciones al equipo de computo.
- c) La administración de costos calculado en horas por semana/mes/año requerido para mantener los sistemas en operación. Algunos sistemas tienen operaciones automatizadas de respaldo total, respaldo parcial, respaldo tipo espejo(mirror), mientras otros necesitan a un administrador de sistemas para supervisar y hacer tareas similares. El costo debe ser calculado con un adecuado rango de horas, para calcular los costos administrativos anuales y de largo plazo.

DIRECCIÓN GENERAL DE BIBLIOTECAS

Basándonos en la tabla1, capacidades de medias de almacenamiento óptico, la comparación de costos debería ser así:

- ¿Cuando, los jukeboxes deben ser requeridos de cada tipo de media (12", 5.25", CD-R y DVD-R) para almacenar 900GB ?
- ¿Cuántos y que tipos de servidores se requieren?
- ¿Que tipo de software se necesita?
- ¿Cuál es el costo de mantenimiento preventivo para cada uno?
- Entre otras cosas.

3.1.4 Opciones de Tecnología WORM.

Con esta tecnología, la información puede ser leída de una media de write-once muchas veces, pero dicha información puede ser escrito en un sector en particular una sola vez. Idealmente, una vez que la información ha sido escrita en el sector de la media, esta no puede ser borrada o sobre-escrita. Con algunas implementaciones, cualquier intento al sobrescribir o rescribir en un área específica da por resultado la destrucción de la información que previamente se encontraba en dicha área, creando un error del sistema. La nueva información debe ser escrita en un área o porción de la media nunca antes usada, y con esto proveer un registro permanente de toda la información escrita en un área particular de la media. De esa forma, la tecnología WORM es ideal para aplicaciones en las que se requiere realizar ensayos de auditoría de todas las transacciones, o que deben demostrar la integridad de ciertos documentos electrónicos. Esto hace que la tecnología WORM sea una opción muy atractiva en las sociedades legales, financieras y de seguros.

3.2 Comparando Costos de Compra vs. Desarrollo de una aplicación completa de digitalización de documentos.

Para tomar la decisión de comprar o desarrollar una aplicación de digitalización de documentos, se necesitan hacer estimaciones razonables para determinar el costo del desarrollo de una aplicación, con la cual podrás comparar posteriormente contra el costo que implica el comprar una aplicación comercial.

Existen cinco fases básicas a considerar al momento de desarrollar una aplicación :

- Codificación
- Pruebas
- Documentación
- Mantenimiento
- Soporte

El costo de cada uno de estos puntos varía dependiendo de cuan compleja se requiera la aplicación, esto en base a los requerimientos. A continuación se explican los pasos para la estimación de costos de los cinco puntos.

3.2.1 Fase de Codificación.

La fase de codificación es requerida para todos los proyectos de programación, generalmente es la que más tiempo consume y la más compleja. El tiempo requerido depende de un numero de factores, los cuales debes de tomar todos ellos en consideración, y estos son los que a continuación se mencionan :

3.2.1.1 Lenguaje de programación.

El lenguaje de programación Visual Basic es el mas común para el desarrollo de proyectos y generalmente se incluyen algunas líneas en lenguaje de programación C o C++. De cualquier manera, en algunos casos, el performance de los requerimientos determina el uso de lenguajes de

programación mas flexibles, y esto da como resultado una adición al programa de trabajo de desarrollo.

3.2.1.2 Conocimiento de la base de datos.

Las aplicaciones de escaneo de documentos son manejadores de bases de datos por naturaleza, y frecuentemente surgen tropiezos por la complejidad, esto por el bloqueo de la base de datos y los problemas de integridad de los datos, especialmente si estos son diseñados para aplicaciones multi-usuarios. Si tu equipo de programación no tiene el suficiente conocimiento en el diseño de bases de datos relacionales y mejoramiento de problemas, estararas en serios problemas con la base de datos y su información durante el desarrollo y mantenimiento de la aplicación.

3.2.1.3 Herramientas de desarrollo APIs (Application Program Interfase).

Esta es una de las áreas más delicadas para la evaluación. Todas las aplicaciones de escaneo requiere el uso de al menos una herramienta de desarrollo y algunas aplicaciones requieren hasta media docena de ellas. Las herramientas de desarrollo APIs son frecuentemente difíciles de aprender y casi siempre presentan problemas de ejecución cuando se esta trabajando en algún punto durante el desarrollo. La evaluación de herramientas y el tiempo que se requiere para aprender los APIs determinan la mayor parte del programa de trabajo para aplicaciones de escaneo.

3.2.1.4 Revisión de errores.

Una completa implementación de revisión de errores puede, en algunos casos, llevarse tanto tiempo como el requerido para escribir el código de programación mismo. La cantidad de revisión de errores que se requiere depende en cuanto gente estará usando la aplicación y que tipo de soporte estará disponible para ellos, por eso es muy importante tener en mente que esto es probablemente el área de diferencia más grande entre aplicaciones comerciales y aplicaciones desarrolladas. Esta es también la parte que más frecuentemente es mal juzgada para la estimación del tiempo de desarrollo.

A continuación se presenta una gráfica que te permite estimar el tiempo de codificación de la aplicación, basada en las características que se planearon para la implementación. Esto asume que la aplicación es desarrollada en lenguaje de programación Visual Basic y que la revisión de errores esta siendo mejorada.

Características	Se necesita herramientas de desarrollo?	Meses Hombre	Es requerido para tu aplicación?	Meses hombre requeridos para tu aplicación
Aplicación básica de escaneo	SI	2		
Separación automática de documentos	NO	1		
Operación en Red (múltiples estaciones de trabajo)	NO	4		
OCR a todo el texto	SI	1		
OCR solamente a una zona del documento	SI	2		
ICR solamente a una zona del documento	SI	1		
Reconocimiento automático de Código de Barras	SI	1		
Asignación automática de la identificación de la forma	SI	1		
Captura a partir de la imagen	NO	2		
Exportación de imágenes a un sistema final	NO	1		
Instalación	SI	1		
Escaneo remoto a través de internet	NO	6		
Revisión de la calidad de de la imagen	NO	2		
Administración centralizada de lotes	NO	2		
Tiempo total de codificación básica				
Sumar 30% si usted requiere un proceso completo en manejo de errores				
Restar 20% si usted no requiere ningún proceso en manejo de errores				
Sumar 20% de la prueba e integración final				
Sumar 10% por la supervisión y administración del desarrollo				
Tiempo Total de Codificación				
Cargar el costo mensual de un programador				
Total global en Costo de Codificación				

Tabla 5 .- Tiempo de codificación de la aplicación.

A continuación se menciona un ejemplo de como usar la gráfica :

- El tiempo total de la codificación básica para una aplicación que requiere los procesos de escaneo, instalación, determinación de una llave de búsqueda, exportación, separación de documentos, y administración por batch o lote es de 9 meses-hombre.
- Para implementar un completo proceso de detección de errores, se deberá sumar 2.7 meses-hombre.
- Sumar 1.8 meses-hombre para cada dispositivo de prueba e integración.
- Sumar .9 meses-hombre por la administración general.

Esta es una aplicación de escaneo moderadamente sofisticado con un total de tiempo de codificación de 14.4 meses-hombre. Un programador de Visual Basic tiene un costo anual de \$90,000, o \$7,500 por mes, por lo tanto el costo total para la fase de codificación da un total de \$108,000

DIRECCIÓN GENERAL DE BIBLIOTECAS

3.2.2 Fase de pruebas.

En general, el tiempo de prueba es calculada mejor como un porcentaje del tiempo estimado de la fase de codificación, y a partir de que considere aplicaciones más complejas, normalmente requiere un mayor numero de procedimientos de prueba, como se ve a continuación.

146948

Características	Es requerido para tu aplicación?	Calculo de Meses Hombre	Meses hombre requeridos para tu aplicación
Pruebas basicas y aseguramiento de calidad		30%	
Pruebas extensas de escanner		20%	
Pruebas de estress		10%	
Pruebas de regresion		20%	
Tiempo total de prueba			
Incluir el costo mensual de usuario dedicado a probar la aplicacion			
Costo total de prueba			

Tabla 6 .- Tiempo requerido para pruebas.

Continuando con el ejemplo anterior :

- El tiempo de pruebas básicas es un 30% del total del tiempo de codificación, o bien, 4.3 meses-hombre.
- Las pruebas importantes toman 1.4 meses-hombre.
- Las pruebas de regresión toman 2.8 meses-hombre.

El tiempo total es alrededor de 8.5 meses-hombre. Tomando en cuenta un ingreso mensual de \$ 5,000 de la persona encargada de hacer pruebas, el costo total de la fase de pruebas es de \$ 42,000.

3.2.3 Fase de documentación.

La documentación puede ser estimada aproximadamente en un 30% del total del tiempo de codificación para un manual técnico.

De cualquier manera, el costo puede variar dramáticamente, dependiendo de cuanta gente se necesita entrenar, el nivel de sofisticación, así también si se elige proporcionar documentación a detalle. A partir de esto utilizamos la siguiente tabla.

Características	Es requerido para tu aplicación?	Calculo de Meses Hombre	Meses hombre requeridos para tu aplicación
Tiempo base de documentación		30%	
Ayuda en línea		15%	
Total del tiempo de documentación			
Incluir el costo mensual de usuario dedicado a la documentar			
Costo total de documentación			

Tabla 7 .- Tiempo requerido para la documentación.

Continuando con nuestro ejemplo, el porcentaje del costo de documentación para nuestro proyecto es de 30% de 14.4 meses-hombre, es decir, 4.3 meses-hombre. Con un cargo al costo de \$5,000 por mes para un documentador técnico, el total del costo para la fase de documentación es de \$21,000.

3.2.4 Fase de mantenimiento.

Los procesos de mantenimiento son requeridos para reparar pequeños detalles, agregar mejoras, y mantener al día los cambios que se requieran por parte de los proveedores de las herramientas que se utilizan para el desarrollo de la aplicación. La siguiente tabla nos muestra como calcular los mantenimientos que se requieren.

Características	Es requerido para tu aplicación?	Calculo de Meses Hombre	Meses hombre requeridos para tu aplicación
Mantenimiento anual basico		15% del tiempo de codificación	
Actualizaciones de sistemas operativos		1	
Actualizaciones a escanners		.5	
Actualizaciones a herramientas para desarrollo de aplicaciones		1	
Actualización a herramientas de compilación		.5	
Tiempo total de mantenimiento			
Incluir el costo mensual de un programador			
Costo total de mantenimiento por año			

Tabla 8.- Estimación de costos de actualización.

Continuando con nuestro ejemplo, tendríamos la siguiente información:

- El tiempo de mantenimiento básico es el 15% del tiempo original de codificación, o 2.2 meses-hombre por año.

- La actualización del sistema operativo, actualización de escanners, y actualización de herramientas de desarrollo, tomaran generalmente cerca de 3 meses-hombre por año.

Por lo tanto, el tiempo total de mantenimiento es por consiguiente 5.2 meses-hombre por año. Como el costo de un programador de lenguaje Visual Basic es de \$7,500, el costo total de la fase de mantenimiento es de \$39,000 por año.

3.2.5 Fase de soporte.

El soporte técnico es requerido generalmente para la capacitación de usuarios, respuestas a preguntas, problemas durante la operación de la aplicación, y mantenimiento al sistema. El total de tiempo requerido depende de cual extenso son las necesidades de tu soporte, como se muestra a continuación en la tabla.

Características	Es requerido para tu aplicación?	Calculo de Meses Hombre	Meses hombre requeridos para tu aplicación
soporte básico de ayuda por teléfono		.5 por usuario	
Capacitación de usuarios		.2 por usuario	
Administración del sistema y mantenimiento		20% del tiempo de codificación	
Soluciones avanzadas y administración de estas		.5 por usuario	
Tiempo total de soporte			
Incluir el costo mensual de una persona de soporte tecnico			
Costo total de soporte por año			

Tabla 9 .- Determinación del tiempo a soporte.

DIRECCIÓN GENERAL DE BIBLIOTECAS

El tiempo de soporte se relaciona principalmente al tamaño del sistema y el numero de usuarios. Continuando con nuestro ejemplo, y asumiendo que el sistema es utilizado por 6 usuarios :

- El tiempo de soporte básico es .5 meses-hombre por usuario, o 3 meses-hombre por año.
- El entrenamiento a un usuario comprende 1.2 meses por año.

- La administración del sistema (el cual depende mas de la complejidad de la aplicación que por el numero de usuarios) es un 20% del tiempo de codificación, o 2.9 meses-hombre por año.

El tiempo total de soporte es por consiguiente 7.1 meses-hombre por año. Tomando en cuenta un costo de \$6,000 por un administrador del sistema, el costo total de la fase de soporte es de \$42,000 por año. (Si es requerido un solucionador de problemas avanzado, el costo deberá incrementarse a 10.1 meses-hombre, o \$60,000 por año.)

3.2.6 Cálculos finales.

El costo total de desarrollo de una aplicación a la medida del cliente, equivale a la suma de los costos de cada fase. Las fases de mantenimiento y soporte necesitan ser calculadas por el tiempo de vida del proyecto, y 5

años es un buen calculo del tiempo de uso de una aplicación. La tabla que a continuación se muestra, nos proporciona el calculo final.

Fase	Costo
Codificación	
Pruebas	
Documentación	
Mantenimiento	
Soporte	
Total Costo Inicial	
Mantenimiento al sistema (2 a 5 años)	
Soporte (2 a 5 años)	
Total Costo por 5 años	

Tabla 9 .- Determinación del costo total del proyecto.

Concluyendo nuestro ejemplo, el costo total inicial del proyecto para el desarrollo de una aplicación es :

- Codificación : \$108,000
- Pruebas : \$ 42,000
- Documentación : \$21,000
- Mantenimiento : \$39,000
- Soporte : \$ 42,000

El costo total inicial del proyecto es de \$252,000. El costo de mantenimiento y soporte por el periodo de 2 a 5 años es de \$ 324,000 , por lo tanto, el costo total por los 5 años de vida del sistema es de \$ 576,000 .

Los costos pueden variar dependiendo de que tipo de aplicación se requiere y que tan extensivo se necesite la documentación y el soporte. Sin embargo, cualquier tipo de aplicación que se necesite, la metodología descrita para el ejemplo de la aplicación, te puede ayudar a estimar tus costos totales y comparar dicho costo contra el costo de una aplicación comercial a comprar, la cual satisfaga tus requerimientos.

4. - INVESTIGACIÓN DE CAMPO.

4.1 Caso Practico.

4.2.1 FEDEX.

Cuando las facturas de mensajería aérea lo son todo, y estas mismas tienen que llegar en la noche del mismo día, o inclusive antes, Fedex esta contando con la tecnología de digitalización para hacer dicho trabajo.

Llevando lo que seguramente es la más grande aplicación en el mundo de digitalización de documentos, en la base de Federal Express localizada en Memphis E.U.A., sé esta integrando un sistema desarrollado de acuerdo a las necesidades de Fedex, en la cual sé podrán ver mas de 900 escanners puestos en estaciones y rampas facilitadas por todo el mundo. Dicho sistema esta acortando el tiempo de procesamiento y captura de datos a cuestión de horas en lugar de días.

"Este es verdaderamente el primer sistema de escaneo que abarca todo el ancho del mundo", menciona Todd Hollenbeck, director administrativo de la división de captura de embarques del corporativo de Fedex. "Nosotros estamos escaneando documentos de mensajería aérea a través del mundo, convirtiendo la información de dichos documentos en información valida para nuestras bases de datos, esto a través de tecnología ICR (Reconocimiento Inteligente de Caracteres), así como también almacenando las imágenes por siete años para que puedan ser consultadas fácilmente en cualquier momento".

Hollenbeck menciona que el sistema podrá manejar un promedio de 1.2 millones de imágenes por día, con un nivel de demanda robusta, esto a partir de la tecnología confiable que se utiliza. La lista de componentes tecnológicos incluyen escanners Bell&Howell, herramientas de procesamiento de imágenes de Kofax, tecnología para la identificación y reconocimiento de documentos de la empresa Top Image Systems, tecnología para la administración de documentos de la compañía Mobius y tecnología para la administración de almacenamiento de la compañía Veritas.

Mas de 5 millones de facturas de mensajería que se manejan diariamente son totalmente automatizadas gracias al sistema de la compañía PowerShip . Aun así, mas de un millón de facturas de mensajería son llenadas a mano cada día. Fedex actualmente envía dichos documentos no automatizados a un proceso central en Memphis, en donde son escaneados y procesadas entre ochocientos mil y un millón de facturas de mensajería. El tiempo promedio del proceso anterior es de tres días, pero en algunas ocasiones puede llevarse hasta mas de seis días por excepciones en el procesamiento. Hollenbeck menciona que el papel principal del sistema distribuido ha probado que el tiempo de proceso ha sido reducido a ocho horas.

El nuevo proceso distribuido comienza con una mezcla de escanners con alimentadores de hojas de cama y automático. Los escanners son configurados con puertos de interface modelo Adrenalina y tecnología de

procesamiento de imágenes VRS (Re-Escaneo Virtual) de la compañía Kofax.

La combinación del puerto de interfase modelo Adrenalina y la tecnología de procesamiento de imágenes VRS es crucial, esto debido a que, los documentos que son llenados con pluma manualmente, en las copias de la factura como tienen un pasante de papel carbón, sale muy débil los detalles de la información escrita. La tecnología VRS empieza con escaneo en escala de grises y corrección automática de inclinación de documentos, corta y mejora el rendimiento bitonal de las imágenes para la identificación de información y su almacenamiento.

Otras etapas fueron tomadas para mantener las cosas simples y libres de errores. Por ejemplo, las herramientas para imágenes de la compañía Kofax, fueron usadas para la identificación del código de barras de la factura, para capturar el número de la factura de mensajería, y para crear una configuración predeterminada para el escaneo con solo oprimir un botón de la aplicación. La compañía Bell&Howell adecuo los escanners de acuerdo a las necesidades requeridas para el escaneo.

Mandy Chubin, director internacional de marketing de la compañía Bell&Howell, comenta, "Nosotros modificamos la configuración manual de los escanners modelo 8000 con el fin de eliminar las configuraciones manuales que pudiera realizar el operador del escanner, así también modificamos los escanners de cama modelo 500FB para habilitar todas las opciones de configuración que te da la tecnología VRS de Kofax".

Fedex planea instalar de uno a dos escanners en las 850 oficinas localizadas en los Estados Unidos de America, así también en cientos de oficinas internacionales con al menos 1,200 escanners en todos ellos. Cada oficina procesara en cualquier parte desde algunos cientos de imágenes, hasta arriba de 8,000 imágenes por día.

Fedex esta usando una aplicación desarrollada a sus necesidades para transferir las imágenes a su oficina central localizada en Memphis. Desde dicho lugar, la tecnología de reconocimiento de la compañía Top Image Systems, es aplicada para reconocer automáticamente los tipos de documentos y extraer los datos incluyendo los números de tarjeta de crédito, así como los nombres localizados en las secciones de envío y recepción, la dirección y los números telefónicos. Las imágenes son administradas por la tecnología desarrollada por la compañía Mobius Management Systems, y estas mismas son almacenadas a través de una aplicación de almacenamiento jerárquico desarrollado por la compañía Veritas.

Hollenbeck menciona, "El verdadero reto fue el manejo de la gran cantidad de imágenes que entraban y salían a diario. Así también, el descifrar el gran volumen de datos, equivalente a 120 Terabytes, los cuales se necesitan tener en línea y continuamente disponibles para su consulta.

En adición a la habilidad de la velocidad del envío de la información de la factura, el sistema de distribución esta dando como resultado una

reducción en el costo de operaciones internacionales, que anteriormente contaba con un alto volumen de envío por fax de facturas de mensajería, así como facturas comerciales, exportación de licencias y otros documentos relacionados en envíos internacionales.

Hollenbeck menciona, "Nosotros eliminamos el envío por fax de cientos de miles de documentos. Ahora nosotros podemos simplemente transmitir imágenes de alta calidad, y estas son traducidas en enormes ahorros en costos de teléfono y mano de obra.

El sistema también puede proveer una información mas completa para el manejo interno del servicio de rastreo. Proporciona un completo manifiesto de los limites de envíos para el mejoramiento de reportes y planeación.

En las fases de planeación que comenzaron desde la primavera de 1999 y aprobado el año pasado, el nuevo sistema de procesamiento de imágenes de Fedex está presente en ocho países en Asia, cuatro países en América Latina y tres países en Europa. Hollenbeck menciona que la compañía espera tener el sistema entero en operación en la siguiente primavera del 2002.

5.- ESTANDARES TECNOLOGICOS DE CLASE MUNDIAL.

5.1 Formatos y tipos de compresión de una imagen.

Los formatos de imágenes TIFF(Compresión Grupo4), JPEG y PDF son los más comunes para la digitalización de documentos, no obstante, los desarrolladores de aplicaciones están buscando mejores alternativas. En una comparación de dos formatos nuevos, DjVu y Lura-Document, con estilos de compresión "wavelet", se crearon archivos dramáticamente más pequeños que se aproximaban al estándar de calidad de la imagen creadas en TIFF, JPG y PDF.

A pesar de los bajos costos en cuanto al almacenamiento de imágenes, el tamaño del archivo continúa siendo de gran importancia en la digitalización de documentos. Las restricciones del ancho de banda son quizá el principal problema para transmitir una imagen, así como descargas

simultáneas de un grupo de imágenes, y en casos donde se manejan imágenes multi-paginas, puede llegar a crear conflictos de tráfico en la red de una empresa o corporativo. El reto del problema aumenta aun más cuando la imagen se necesita acceder vía Internet o vía correo electrónico.

Los formatos de imágenes DjVu y LuraDocument usan combinaciones similares de segmentación de imágenes y técnicas "wavelet". Dicha técnica se basa en segmentaciones, las cuales guardan un alto grado de concentración de información de texto en una capa separada. La compresión

Wavelet descarta para el resto de las áreas de la imagen, bits que no son necesarios para preservar una calidad perceptible de la imagen.

Mientras ambos formatos pierden una pequeña parte de la información de la imagen original, así también las áreas fuertes como el texto también pueden perder un poco de la nitidez. Los ratios de compresión de la tecnología wavelet pueden exceder de 100 a 1, mientras que la compresión del formato JPEG no excede más de 30 a 1 sin degradar la visibilidad de la imagen.

En un principio, los formatos DjVu y LuraDocument eran similares. Ahora, dichos formatos usan diferentes algoritmos y métodos de codificación propietaria. Cada formato requiere su propio plug-in para ver las imágenes en el navegador de Internet.

Originalmente desarrollado por AT&T, DjVu es ahora comercializado por Lizard Tech, la misma compañía que promueve el formato de compresión para imágenes gráficas llamado Mr.Sid . Lizard Tech esta promoviendo actualmente el formato de imagen DjVu con una campaña de mercadotecnia agresiva en el mercado del usuario final, con los productos de escaneo y conversión de imágenes llamado Solo. Estas herramientas ofrecen además tecnología OCR (Reconocimiento Óptico de Caracteres), la cual te permite crear un estilo de Imagen-texto, en el cual te permite buscar texto en la imagen misma.

Para poder ver la imagen en formato DjVu, Lizard Tech ofrece un plug-in gratis, el cual puede ser bajado del sitio de Internet de esta compañía (www.lizardtech.com).

La compañía menciona que actualmente existen 25 millones de exploradores de Internet con la capacidad de ver las imágenes en formato DjVu. Además, un control ActiveX a sido sacado al mercado, el cual permite ver las imágenes con formato DjVu desde cualquier aplicación Office de Microsoft.

El formato de imagen LuraDocument se deriva del desarrollo del formato de imágenes gráficas llamada LuraWave, desarrollado por Lura Tech. Esta compañía es la que más éxito ha tenido en Europa, donde muchas compañías de administración de documentos a adoptado este producto. Esta tecnología esta siendo comercializada para desarrolladores de aplicación en Norte América por la compañía Menlo Park.

DIRECCIÓN GENERAL DE BIBLIOTECAS

5.1.1 Comparación de formatos de imágenes convencionales.

Las pruebas que a continuación se presentan, apuntan a formatos de imágenes que comúnmente se usan; el propósito fue el escanear documentación de negocios, así como documentación de mercadotecnia. Se escanearon un cierto volumen de documentos, a una resolución de 200dpi (puntos por pulgada), y guardados en los formatos TIFF grupo4, PDF, JPEG, DjVu y LDF.

Tipos de Documentos (Escaneado a 200dpi bitonal)	TIFF Grupo 4	JPEG	PDF Tipo Imagen	PDF Tipo Imagen+Texto	PDF Normal	DjVu Tipo- Imagen	DjVu Tipo Imagen+Texto	LuraDoc
Aplicacion Contable	91K		101K	113K		34K	39K	57K
Recibo de Aerolinea	17K		21K	26K		9K	10K	
Cheque	25K		30K	33K		19K	19K	
Libro de Jardineria	47K		54K	63K	68K	28K	29K	
Documento Multipagina	105K		136K	159K		26K	32K	61K
Ticket de Estacionamient	33K		39K	44K		22K	22K	31K
Recibo Telefonico	15K		22K	26K		6K	8K	
Periodico	43K		54K	64K	10K	10K	13K	
Tipos de Documentos (Escaneado a 200dpi, 8- bit en escala de grises)	TIFF Grupo 4	JPEG	PDF Tipo Imagen	PDF Tipo Imagen+Texto	PDF Normal	DjVu Tipo- Imagen	DjVu Tipo Imagen+Texto	LuraDoc
Cheque	1,258K		227K	228K		30K	31K	
Ticket de Estacionamient	1,243K		270K	270K		35K	36K	
Tipos de Documentos (Escaneado a 200dpi, 24-bit color)	TIFF LZW	JPEG	(*1)PDF con buscador de texto	(*2)PDF con buscador de texto	PDF formateado Texto+Grafica	DjVu Tipo- Imagen	DjVu Tipo Imagen+Texto	LuraDoc
Aplicacion Contable	5,224K	1,452K	1,488K	1,299K	396K	43K	47K	80K
Anuncio	11,586K	2,662K	2,777K	2,380K	2,331K	163K	168K	192K
Recibo de Aerolinea	2,814K	776K	720K	757K		27K	28K	30K
Imagenes de cheques	2,934K	766K	755K	713K		40K	40K	36K
Libro de Jardineria	4,765K	1,199K	1,258K	989K	695K	61K	64K	81K
Forma de Seguro	9,263K	2,162K	1,862K	1,764K		76K	87K	101K
Ticket de Estacionamient	3,720K	895K	856K	808K		49K	52K	55K
Declaracion de Utilidades	3,732K	1,024K	1,049K	903K	215K	31K	33K	41K

(*1) Usando compresion de color "High quality JPEG"

(*2) Usando compresion de color "ZIP"

Tabla 10.- Comparación de formatos de imagenes.

Como podemos apreciar en la gráfica, DjVu y LuraDocument nos dieron archivos dramáticamente más pequeños. El formato Bitonal TIFF Grupo4, PDF en formato imagen solamente y PDF en formato imagen mas texto, fueron cerca de 3 veces más grandes, en promedio, que el formato de imagen DjVu. Aunque el uso del formato LuraDocument no es para uso bitonal, algunos ejemplos de imagen fueron cerca de 50% más pequeños que las imágenes en formato TIFF y PDF.

Los resultados de la comparación fueron drásticos en las imágenes de color, teniendo como resultado que la imagen con formato TIFF pero con compresión LZW fue 85 veces más grande que la imagen en formato DjVu. Las imágenes en formato LuraDocument fueron también pequeñas, sin embargo, las imágenes en formato DjVu fueron cerca de 15% más pequeñas en promedio.

El formato normal de PDF es algunas veces mencionado como la mejor alternativa de compactación, sin embargo, este formato convierte las imágenes a documentos electrónicos. Con el formato PDF no se puede manejar un registro de imágenes.

Al darnos el tipo de compresión wavelet en las imágenes de un tamaño mucho más pequeño, tu puedes esperar también un compromiso en la calidad de la imagen. El mayor reto para este tipo de compresión son las copias de los tickets de estacionamiento y los documentos que son difíciles de leer, incluso del documento original, las cuales solamente pueden ser leídas si se guarda la imagen a color o en escala de grises.

En estas imágenes, las imágenes con tipo de compresión wavelet no pudieron ser elidas como las imágenes con formato TIFF, JPEG y PDF. Las imágenes en formato DjVu retuvieron ligeramente mas de la información importante a comparación de las imágenes con el formato LuraDocument, sin embargo, el suave mejoramiento que le da al texto con una débil letra escrita con pluma, hizo ver dicho texto de la imagen difuso. El mismo texto fue cortado en las imágenes con formato LuraDocument, e incluso, se pierde

el texto escrito muy fino, a mano y con pluma. Así también, las imágenes con formato LuraDocument sufrieron algunos detalles con las partes del documento donde le da un poco de brillo a las imágenes cuando está con sombra y letras, puesto que al darle brillo, extrae la sombra y las letras, debiendo de conservar las letras en sí.

En base a los resultados de este estudio, las compañías que comercializan los formatos DjVu y LuraDocument, mencionaron que dichas situaciones pueden arreglarse desarrollando una aplicación especial, esto mediante las herramientas que ofrecen para el desarrollo de aplicaciones a la medida del cliente. Con dichas herramientas, la persona encargada de desarrollar la aplicación puede hacer varias pruebas, ajustando las opciones de la configuración de mejoramiento de la imagen (en secciones con sombra o con brillo) de documentos que son escritos con pluma y con un texto muy fino o débil.

Si las necesidades de digitalización se basan en documentos sin sombras donde se encuentra el texto, y con texto escrito con una calidad aceptable o texto impreso, los formatos DjVu y LuraDocument pueden ser la mejor solución. Pero si tus documentos tienen texto escrito a mano de poca calidad, o existen áreas donde escribes en un área con sombra, entonces los formatos DjVu y LuraDocument no son un formato conveniente para documentos con dichas características.

Los formatos DjVu y LuraDoc soportan imágenes multipáginas, lo cual es una ventaja sobre el formato JPEG, pero ninguna de los tres formatos mencionados no soportan la capacidad de etiquetado de imagen que están disponibles en los formatos TIFF y PDF. Así también, los formatos DjVu y LuraDocuments carecen de opciones estándar de exportación para su manejo en los sistemas de administración de imágenes más comunes.

El futuro se balanza en la promesa de un formato final estándar JPEG 2000/JPM, el cual alcanzara una alta compactación del documento, basado en un tipo de compresión wavelet específico para digitalización de documentos. Este futuro formato estándar también podrá ser soportado por los navegadores de Internet. Inclusive, si esto llega a liberarse en un año mas, se necesitara mas tiempo mientras se actualizan las versiones de los navegadores de Internet y plug-ins. Hasta ese entonces, las alternativas de formatos de imágenes están disponibles en la actualidad para quien necesite utilizarlas.

DIRECCIÓN GENERAL DE BIBLIOTECAS

6.- CONCLUSION.

Como pudimos apreciar en los capítulos anteriores, existen una amplia gama de factores críticos para obtener una solución de digitalización a documentos, los cuales, si no se selecciona el adecuado en cada una de las fases del proceso, se podría terminar en un fracaso total, y por ende, en pérdidas de tiempo, recursos materiales y recursos económicos.

A partir de lo anterior, es determinante elaborar una serie de cuestionamientos, estos durante la planeación de la solución, para de esa manera reducir al máximo cualquier tipo de riesgo que pudiese suscitarse durante el proceso de digitalización.

Para evitar al máximo cualquier tipo de problema en el proceso de digitalización, a continuación se mencionaran los pasos que se deben seguir,

así como el curso a tomar en cada uno de estos, para poder tener una operación y administración de imágenes exitosa.

DIRECCIÓN GENERAL DE BIBLIOTECAS

1.- Definir la necesidad específica de la compañía.

En primer lugar debemos de preguntarnos “¿Qué es lo que queremos?”, es decir, solamente reducir el espacio físico del archivo convencional a su mínima expresión?, así también, ¿Se requiere explotar la información de los mismos documentos como herramienta potencial en la empresa?.

2.- Analizar el estado físico y clasificación de los documentos.

Teniendo identificada la necesidad específica de la compañía, será necesario analizar y determinar el estado físico y la clasificación de los documentos que se requieren escanear, puesto que a partir de esto se determinará si es viable la posibilidad de escanear dichos documentos, así también la identificación de cada uno de estos.

3.- Definición de clave única de cada documento-imagen.

Una vez que se haya realizado el análisis del estado y clasificación de los documentos, es necesario determinar una clave única o nombre único que identificará el origen de este mismo, con el fin de determinar las siguientes características:

- Tipo de documento.
- Departamento de donde proviene.
- Persona o empresa a la que pertenece.
- Fecha en que se escaneó el documento.
- Consecutivo del documento, esto para identificar si es una sola hoja, o varias de un tipo o departamento o
- Entre otras características personalizadas para cada empresa.

Para definir la clave única del documento, es importante definir un control de grupos o divisiones de identificación, especificando el número de grupos que la empresa requiera para determinar una clave única para cada

imagen. Por lo general se recomienda un mínimo de 2 grupos, además del consecutivo.

Como ejemplo de una clave única del documento, se tiene el siguiente nombre de imagen :

FARHEMPR001.tif

En donde :

- Las letras FA identifica que es una factura.
- Las letras RH identifica que proviene del departamento de recursos humanos.
- Las letras EMPR identifica las siglas de al EMPRESA.
- Los numeros 001 identifica que es la primer imagen de la clasificacion FARHEMPR
- .tif es la extension del formato de la imagen.

4.- Definir el tipo de formato de la imagen.

Al tener ya definido el control y nomenclatura de los documentos a escanear, se debe determinar el tipo de formato y compresión de la imagen. En la actualidad, el formato definido por la AIIM como estandar para las imagenes, es el formato "TIF" con compresión "GROUP IV", ya que es un formato que se utiliza en casi la totalidad del software de consulta de imagenes. Aunque pueden ser utilizados los formatos DjVu o LuraDocument con el objetivo de tener un menor tamaño de imagen, se necesita de tecnología propietaria de las compañías que tienen los derechos de uso de dichos formatos para poder consultar estas mismas.

5.- Definir los tipos de acceso.

Al escanear los documentos, se deberán almacenar las imágenes en dispositivos de almacenamiento, esto de acuerdo a la frecuencia con la que se requiere consultar, así como también el volumen que se requiere tener en este mismo. Como un ejemplo de este punto se podría definir como tipo de dispositivo un CD-R para el almacenamiento, teniendo dos opciones : consultarlo directamente mediante CD-R insertando en forma manual el CD-R al lector individual, o bien, en caso de que el volumen de información sobrepase 1 TeraByte, se recomienda usar un JukeBox para el manejo automático de los CD-R. Cabe mencionar que dicha definición depende también del presupuesto que se tenga disponible para el proyecto.

6.- Determinar la compra o desarrollo de la aplicación a utilizar para el escaneo de los documentos.

Ya que se tiene definido que se requiere, que control se va a utilizar, que formato y compresión de imagen, así como los tipos de acceso, será necesario realizar un análisis para determinar si se compra o desarrolla la aplicación a utilizar para el escaneo de los documentos.

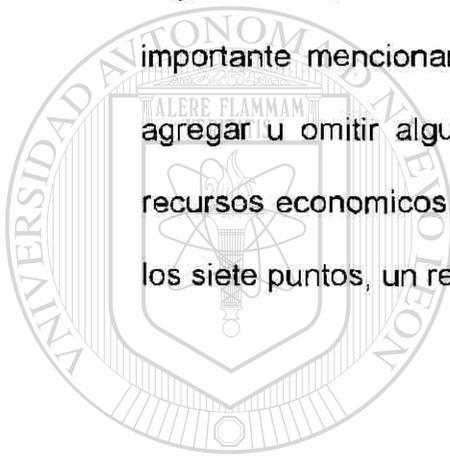
7.- Seguridad.

Para incluir en los planes de contingencia, como seguridad, es indispensable realizar cuando menos una copia de las imágenes que se están generando, con el fin de poder emplearlas como reemplazo en caso de una situación de desastre. Así también, se deberá almacenar este mismo en un lugar distante al que se tiene la información en línea.

7.- Control.

Es importante llevar un control de la información que se esta consultando, esto en los casos en que se haya definido tener imagenes recientes para consulta en línea (Near-line) , asi como tener imagenes de cierta antigüedad en archivo (Off-line).

El cumplimiento de los puntos anteriores es fundamental para un seguimiento y control exitoso de un proceso de escaneo. No obstante, es importante mencionar que cualquier empresa esta en todo el derecho de agregar u omitir algun punto, esto en base a el criterio, circunstancias y recursos economicos de esta misma, resultando de la posible alteración de los siete puntos, un resultado incierto, ya sea positivo o negativo.



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

®

DIRECCIÓN GENERAL DE BIBLIOTECAS

BIBLIOGRAFIA

- Libros

a) Manual del escanner para profesionales

Mc Graw-Hill

Emil Ihrig y Sybil Ihrig

b) CDIA Test Preparation Guide

The Rehinner Group

- Revistas

a) e-doc magazine

b) Imaging and Document Solutions

- Internet

a) www.edocmagazine.com

b) www.techinfocenter.com

c) www.imrgold.com

d) www.isit.com

e) www.imation.com

f) www.fujitsu.com

g) www.documentimaging.com

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

®

DIRECCIÓN GENERAL DE BIBLIOTECAS

APENDICE

A. Caso de la Embajada de Estados Unidos de Norteamérica en México : Solución al manejo de documentación del proceso de la Visa.

El tiempo de proceso para el manejo de la papelería y fotos en la embajada de Estados Unidos en México, fue volviéndose mas lento para los miles de tramites diarios que se realizan para la expedición de la visa. Cada vez mas líneas de gente con su papelería e identificación con foto se incrementan buscando obtener la Visa.

Reconociendo la perdida de recursos y tiempo, el personal de la Embajada de Estados Unidos en México analizó la necesidad de hacer un cambio

revolucionario a dicho proceso. Implementando una solución de digitalización de documentos predesarrollada de la compañía Kofax, la embajada estableció un dinámico y grandioso proceso desarrollado a prueba de fraudes.

1 El problema.

En la embajada de los Estados Unidos de América con sede en la ciudad de México, cada día asisten de dos mil a tres mil personas para el tramite de la visa. Cada persona es entrevistada y fotografiada, creando con esto miles de documentos originales cada día. Debido a la complejidad de las leyes

concernientes a las visas, muchas solicitudes son rechazadas diariamente, requiriendo con esto que las personas regresen a la embajada para volver a procesar su solicitud de visa. Para estos casos, el personal de la embajada debe de acceder la documentación de la solicitud previa para confirmar la identidad de la persona.

Hasta hace poco, la identificación por foto fue incluida en la solicitud de la visa, y la información es capturada en una base de datos maestra. Sin embargo, los documentos escritos a mano y fotos son puestos en paquetes y son guardados en archiveros.

Cuando una persona vuelve por un reproceso de la visa, la información de su solicitud es accesada a través de archivos electrónicos. En la entrevista con el oficial, se le pide a la persona interesada una identificación con foto para confirmar su identidad. Es aquí donde el proceso cae en un bache. Con

la lenta operación del personal de archivo, a través de archiveros con cajones en desorden, cada foto o papel de la solicitud toma entre diez o veinte minutos en localizarlos. El estancamiento es alto. Los oficiales de inmigración responsables de la revisión y aprobación de miles de solicitudes se encargan de tomar decisiones que envuelven cualquier tipo de cuestionamiento legal. En estos casos, es imperativo que la identificación misma sea documentada para asegurar cualquier tipo de fraude. Aunque el tiempo que se tomó en el proceso de confirmación de información fue necesario, los problemas de almacenamiento y consulta fueron creciendo, así como también la frustración por parte del personal de la embajada y los solicitantes.

2 La solución.

Jim Sundstrom de la embajada de los Estados Unidos de Norteamérica reconoció que necesitaba guardar la información y fotos del solicitante, todo junto, bajo un sistema, el cual ellos pudieran acceder electrónicamente para una rápida y segura identificación. Los oficiales buscaron un método de escaneo para fotos, así como también para documentos escritos en forma manual con pluma, y almacenar estos mismos en conjunto, para una fácil consulta.

“Todo lo que tenía era una idea”, comentó Sunstorm. “Al principio esto se veía como una tarea imposible. El uso de fotos a color es un factor primordial para la identificación, y nada se veía apto para el escaneo a color en un ambiente de producción.”

Sundstrom buscó sugerencias de expertos en tecnología que pudieran ser aptos para proveer la solución. Muchas compañías le comentaron que lo que el requería simplemente no puede ser realizado. Entonces ellos contactaron a Thomas Hinkle de la compañía data Management Internationale. Hinkle, distribuidor de aplicaciones de Kofax y Kodak, trabajó con la embajada de los Estados Unidos de Norteamérica para diseñar e implementar una solución altamente exitosa. Se conjuntó la aplicación de Kofax Ascent Capture y el scanner de color Kodak de alta velocidad para escanear y almacenar las solicitudes doble lado con fotos para tener un acceso electrónico. Ahora, las solicitudes son fácilmente consultadas por nombre,

fecha o número de identificación en segundos, y con una disponibilidad instantánea de fotos, información escrita manualmente con pluma y firmas.

“Hemos ahorrado de tres a cinco horas de trabajo por día en los procesos de consulta”, comentó Sundstrom. “La solución de escaneo Ascent Capture es una poderosa herramienta anti-fraudes, así también agilizamos el proceso de toma de decisiones en el que los oficiales están involucrados.”

3 Un modelo de solución.

“El éxito en nuestra solución atrajo la atención de otras embajadas en toda Latinoamérica,” comentó Sundstrom. “Nosotros establecimos una junta en la ciudad de México para todos los oficiales de los consulados de Latinoamérica para que observaran el proceso directamente”.

“Ellos estuvieron impresionados por la facilidad de uso de la aplicación Ascent Capture, así como la escalabilidad y adaptabilidad a otras

aplicaciones, con diferentes lenguajes y dialectos,” comentó. “Esta tecnología ha sido muy fácil de enseñar a nuestro personal, además, se tiene una operación eficiente y libre de errores.”

B. Caso de la empresa Black&Decker aparatos eléctricos - almacena mas que taladros y cortadoras de césped .

1. La compañía.

Black and Decker, localizada en Gothenburg, Suiza, es proveedor líder en el mercado de aparatos y herramientas para el hogar en toda el área nórdica. Actualmente la compañía produce 3,000 facturas semanalmente a minoristas, los cuales venden sus productos.

2. El reto.

Debido a la gran cantidad de facturas producidas cada semana, la búsqueda o rastreo a facturas de clientes ha venido incrementando el tiempo empleado para este mismo, debido a su método actual de almacenamiento de documentos. Las facturas son almacenadas en archiveros físicos

ordinarios, los cuales iban creciendo en cantidad cada semana.

Black and Decker evaluó su situación y llegó a la conclusión que ellos simplemente necesitaban una mejor manera para almacenar las facturas de sus clientes, haciendo con esto una mejor manera de obtener las facturas cuando se necesiten. Ellos determinaron que utilizando un dispositivo CD-ROM era una buena opción. Querían un sistema que les permitiera a ellos seguir trabajando con su actual red Novell, y de esa manera seguir con el personal actual, así como un sistema con excelentes capacidades de búsqueda y recuperación.

3. La solución.

Black and Decker encontró que la aplicación Alchemy fue la aplicación para dispositivos CD-R que ellos necesitaban para resolver completamente sus problemas de almacenamiento de facturas, y esto con un pequeño cambio en su sistema actual. Con la aplicación Alchemy, ellos ahora pueden archivar sus facturas en CDs, haciendo con esto un proceso total de almacenamiento de facturas más efectivo. En lugar de buscar las facturas en sus archivos físicos, el personal de Black and Decker busca estas mismas en CD.

En adición a esto, el departamento de mercadotecnia ahora esta manejando campañas de soporte especial a productos por búsqueda de numero de cliente y numero de producto de las facturas generadas durante el periodo de promoción. Cuando Black and Decker empieza una campaña de publicidad para un producto en particular, la mercancía es entregada a tiendas de menudeo y estas son facturadas a un precio de campaña especial del producto. Después de la promoción, las tiendas de menudeo pueden devolver a Black and Decker la mercancía no vendida.

Antes de usar la aplicación Alchemy, los empleados tenían que buscar manualmente las facturas emitidas durante el tiempo de campaña para verificar el pago del cliente. Ahora con la actual aplicación de almacenamiento se pueden hacer búsquedas de facturas de cientos de tiendas que se localizan en toda el área Nórdica. Los empleados tienen que extraer las facturas al tiempo en el que la campaña empieza, así como también buscar un producto en particular que fue entregado. Esto llevaba

horas, retrasando el proceso de emisión de crédito a las tiendas. Ahora, una vez que la mercancía es regresada, los empleados usan la aplicación Alchemy para buscar mediante el número de cliente o tipo de producto. En segundos, el archivo correcto del cliente es encontrado y el crédito puede ser emitido en ese mismo momento.

4. Retorno de la inversión.

De acuerdo con Alf Karlson, Administrador de sistemas en Black and Decker, "El tiempo que estamos ahorrando es fenomenal."

La aplicación Alchemy es la de mayor uso en Black and Decker, ahorrando tiempo, así como también manteniendo contacto con importantes tiendas de menudeo.

5. Resumen de beneficios.

- a) Acortamiento del ciclo para la emisión de créditos
- b) Mejoramiento en el servicio a clientes
- c) Actualización no costosa del sistema
- d) Reducción del espacio del almacenamiento del archivo

