

Capítulo 1 – Antecedentes del tema

1.1 Dato, Información y Conocimiento

El dato es una representación simbólica de un hecho, como texto, números, gráficos, imágenes, sonido o video, el dato no tiene valor semántico por sí mismo.

La información es un dato en contexto, podemos interpretar el dato con base en el contexto que rodea al dato, el contexto incluye:

1. El significado del negocio y los términos relacionados al dato.
2. El formato en el cual el dato es presentado.
3. El plazo representado por el dato.
4. La relevancia del dato para quién lo usa.

Los datos son la materia prima para que crear información, y el resultado de la información guía nuestras decisiones es decir que la información contribuye al conocimiento. El conocimiento es entonces la comprensión, sensibilización, conocimiento y el reconocimiento de una situación y la familiaridad con su complejidad. El conocimiento es información en perspectiva integrada en un punto de vista basada en el reconocimiento y la interpretación de patrones tales como las tendencias, formado con otro tipo de información y experiencia. A continuación en la Figura 1. 1 se muestra la relación entre Dato, Información y Conocimiento.



Figura 1. 1 Dato, Información y Conocimiento

Los datos dentro de un contexto generan información, ésta sumada a una perspectiva genera conocimiento.

Fuente Imagen: (Solís, 2010)

1.2 Información no estructurada

Es cualquier documento, archivo, gráfica, imagen, texto, reporte, forma, video, sonido grabado, etc. que no ha sido etiquetado o estructurado de alguna manera en registros o filas y columnas.

Los datos no estructurados se encuentran en diferentes tipos de formatos electrónicos, como se muestra en la Figura 1. 2, incluyendo procesamiento de texto, mail electrónico, archivos planos, hoja de cálculo, archivos de XML, mensajes transaccionales, reportes, gráficas de negocio, imágenes digitales, microfichas, grabaciones de video y audio. Y una enorme cantidad de datos no estructurados se encuentran en archivos de papel.



Figura 1. 2 Contenido estructurado y no estructurado

El contenido generado de una empresa comprende actualmente 20% información estructurada (filas y columnas) y 80% información no estructurada (imágenes, páginas web, correo electrónico, audio y video, etc.)

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

1.3 Administración de Contenido Empresarial

Content Management o también conocida como Enterprise Content Management (ECM), en español, la Gestión de Contenido Empresarial es una estrategia transaccional tomada principalmente de las empresas de tecnologías de la información para tener el control sobre captura, almacenamiento, acceso y uso de datos de la información almacenada fuera de la base de datos relacional, es decir de la información no estructurada, mediante procesos, técnicas y tecnologías.

De cierta manera es equivalente a la gestión de las operaciones para los datos de una base de datos relacional (información estructurada), ya que se pretende que la mayoría de los datos no estructurados tengan una relación directa con los datos almacenados en las bases de datos relacionales.

Se enfoca en la integridad y el acceso de la información, resultando en una efectiva recuperación y reutilización de la misma.

De manera general, la Gestión de Contenido Empresarial observa dentro de cada archivo y trata de identificar y usar los conceptos contenidos dentro de éste, para clasificar y así poder gestionar a través de procesos, técnicas y tecnologías.

De acuerdo a la Figura 1. 3 y como parte de una solución completa para una empresa, la Administración de contenido Empresarial debe gestionar la Taxonomía, la clasificación, el modelo de objetos, la estructura de arquitectura, ILM (Information Lifecycle Management), integridad, estándares técnicos, dinamismo y seguridad.

Administración de Contenido

- **Definición:** Planear, implementar y controlar actividades para almacenar, proteger y acceder a los datos encontrados dentro de los archivos electrónicos y archivos físicos (incluyendo texto, gráficos, imágenes audio y video)
- **Objetivos:**
 1. Salvaguardar y asegurar la disponibilidad del valor de los datos almacenados en formatos poco estructurados.
 2. Habilitar la efectiva y eficiente recuperación y el uso de datos e información en formatos no estructurados.
 3. El cumplimiento de las obligaciones legales y las expectativas del cliente.
 4. Asegurar la continuidad del negocio a través de la retención, recuperación y conversión.
 5. El control del almacenamiento documental con costos operativos.

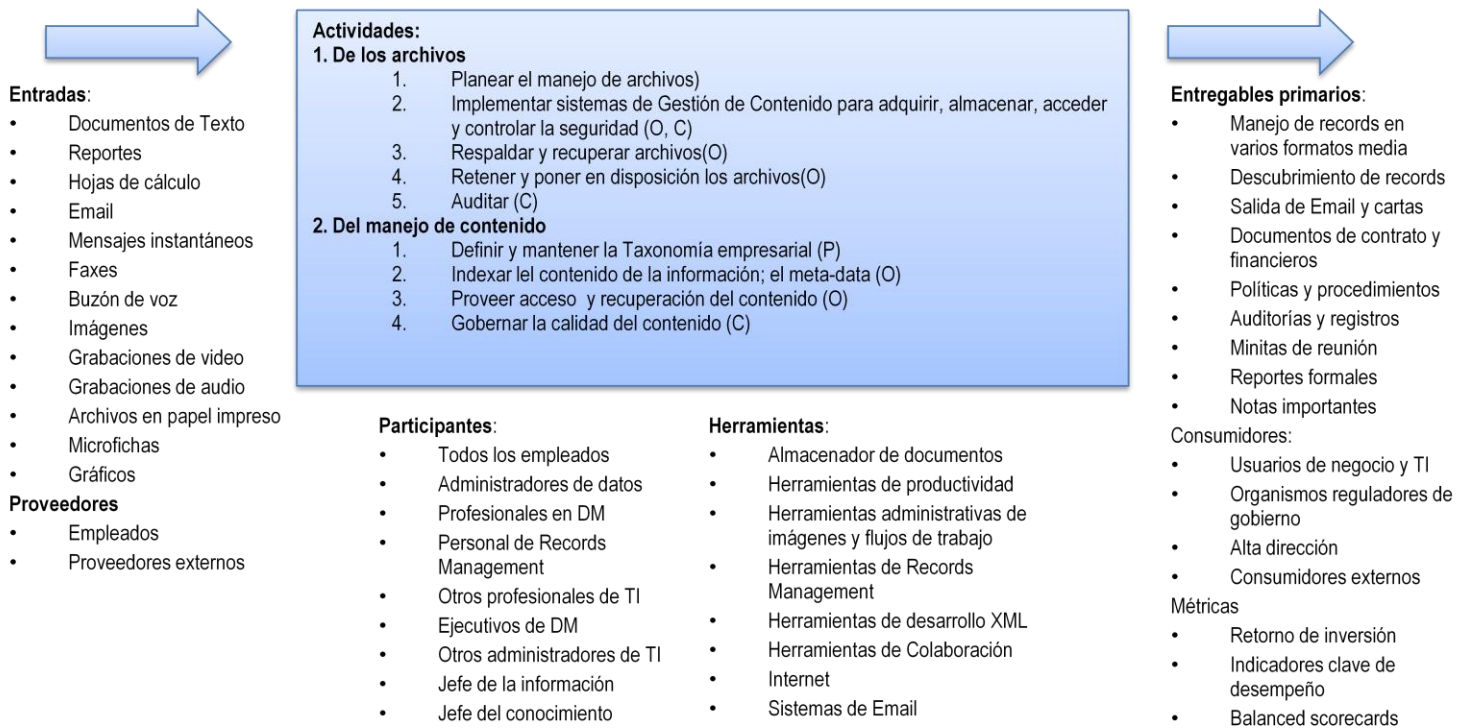


Figura 1. 3 Diagrama de Administración de Contenido

La figura muestra a la Administración de contenido como un proceso; con definición establecida del proceso y objetivos que debe cumplir utilizando ciertas entrada las cuales forman parte de la materia prima que se procesará con ciertas actividades en este caso de los archivos y del manejo del contenido, para llevar a cabo las actividades se nombran a los participantes y las herramientas necesarias y finalmente se consolidan los entregables primarios enunciando los consumidores y métricas.

Fuente Imagen: (Mosley, M. y Brackett, M., 2009)

1.3.1 Taxonomía

Habitualmente el término taxonomía se emplea para ordenar a los organismos vivos en un sistema de clasificación compuesto por taxones anidados, donde cada taxón es un grupo de organismos emparentados asignándoles un nombre, descripción y tipo, es así como se emplea para ordenar todo el contenido que va a ser almacenado en el sistema de almacenamiento, es decir que se ordena en taxones el contenido que tenga relación entre sí, por su contenido, una vez agrupados se les asigna un nombre al grupo de contenido, una descripción detallando el contenido que poseen en común, y finalmente un tipo de contenido. La Figura 1. 4 ejemplifica cómo se organiza una taxonomía.

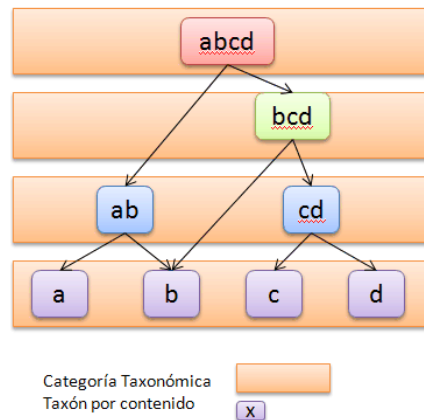


Figura 1. 4 Taxonomía de contenido

Un conjunto de archivos de contenido (abcd) posee características de contenido que permiten clasificar a cada uno de sus elementos dentro de una categoría y sub categoría (bcd, ab, cd, a, b, c y d).

Fuente Imagen: (Solís, 2010)

1.3.2 Catalogación

Es la relación ordenada de elementos que pertenecen a un mismo conjunto, en términos de contenido, la catalogación es una relación de contenido distinguido por las características que éste posee, es decir los metadatos y tipología documental, ver Figura 1. 5

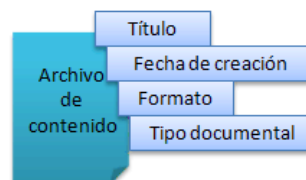


Figura 1. 5 Catalogación de contenido

Un archivo de contenido (imagen, video, documento, etc.) se cataloga a través de las características que lo describen como archivo electrónico.

Fuente Imagen: (Solís, 2010)

1.3.4 Modelo de Objetos

El modelo de objetos se diseña para permitir la administración de cualquier contenido relacionado a sus procesos de tal manera que hay que modelar los objetos que serán administrados durante el proceso de negocio, un objeto es un componente del catálogo (explicado anteriormente) que posee instrucciones que lo administran. La Figura 1. 6 muestra cómo del catálogo de contenido se agregan las instrucciones que le serán aplicadas al contenido.

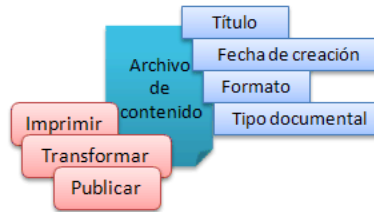


Figura 1. 6 Modelo de Objetos

Un archivo de contenido (imagen, video, documento, etc.) se modela como un objeto a través de las instrucciones que se le deben aplicar durante su tiempo de vida.

Fuente Imagen: (Solís, 2010)

1.3.5 Information Lifecycle Management (ILM)

Information Lifecycle Management (Gestión del ciclo de vida), es un concepto aplicado al contenido para poder controlar el ciclo de vida de los archivos, un archivo de contenido es creado o adquirido, almacenado y mantenido, usado y eventualmente destruido, en el transcurso de esta vida, el contenido puede ser extraído, exportado, importado, validado, editado, actualizado, limpiado, transformado, convertido, integrado, segregado, agregado, referenciado, revisado, reportado, analizado, respaldado, recuperado, archivado y retirado antes de que eventualmente sea borrado, todos los estados del ciclo de vida del contenido están asociados a costos y riesgos dentro del negocio, por esta razón es importante diseñar un ciclo de vida adecuado a cada tipo de contenido En la Figura 1. 7 se muestra un ejemplo del ciclo de vida.



Figura 1. 7 Information Lifecycle Management

Ejemplo de un ciclo de vida, desde la creación del archivo de contenido hasta su retirada, el ciclo de vida puede tener estados en los cuales se modifique, revise, apruebe y publique, teniendo la posibilidad de regresar en caso de que por ejemplo no se apruebe, al estado Revisar.

Fuente Imagen: (EMC, Hoja de Datos h2810.2, 2008)

1.3.6 Procesos

La empresa siempre se encuentra organizada con base en distintos procesos que definen las actividades que se deben realizar, como parte del apoyo al negocio, dentro de ellos existen procesos administrativos que tienen relación con el contenido. La administración de contenido empresarial se encarga de diseñar los procesos que se deben llevar a cabo para la gestión del contenido tomando en cuenta las reglas de negocio que lo afectan.

1.3.7 Seguridad

El manejo de la seguridad del contenido abarca la planeación, el desarrollo y la ejecución de las políticas de seguridad y procedimientos que proveen la autenticación, autorización, acceso y auditoría al contenido. Las políticas de seguridad y procedimientos aseguran que los usuarios correctos interactúen con el contenido en el momento preciso y restringir el acceso a usuarios y tiempos inapropiados.

1.3.8 Integral

La administración de contenido empresarial debe tener un enfoque integral con respecto a:

- El diseño de todos los conceptos anteriores, debe de constituirse como uno solo, es decir que los estados del ciclo de vida se relacionan con el modelo de objetos, con la catalogación, la taxonomía y los procesos.
- La evolución del negocio, la empresa es un ente vivo, que consta de capital humano, procesos, tecnología e información, todos ellos cambian constantemente, el objetivo de la integridad debe permitir la adaptación del diseño de todos los conceptos anteriores para que aún pueda gestionar el contenido, esto no significa que sea estático, pues como parte del ente también debe evolucionar, pero el cambio no debe ser tan drástico si se realiza un diseño adecuado.

1.3.9 Estructura de Arquitectura

El Contenido Empresarial tiene el apoyo de distintas herramientas tecnológicas que facilitan la automatización y administración del contenido, cada una de ellas posee características distintas para construir la infraestructura en Hardware y Software para poder llevar a cabo la gestión de contenido. Y que además se tiene que ajustar a las necesidades del negocio.

1.3.10 Estándares técnicos

Como ya se ha mencionado con anterioridad, las herramientas tecnológicas poseen características distintas entre ellas para poder construir la infraestructura, en este caso, es necesario conocer los límites de la misma, porque todas se pueden personalizar en distintos niveles ya que no son plataformas que impidan hacer modificaciones en el código, pero muchas veces resulta contraproducente hacer tantas modificaciones como uno quiera ya que existen riesgos de funcionamiento, tiempo de implementación y económicos. Se deben establecer los estándares técnicos de cada plataforma y limitar las personalizaciones para asegurar el éxito de la implementación.

1.4 Sistema de Administración de Contenido Empresarial

Un Sistema de Administración de Contenido Empresarial es la tecnología que provee una solución sistémica para que la Gestión de Contenido pueda llevarse a cabo, permite capturar, organizar, almacenar y distribuir contenido no estructurado dentro y fuera de la empresa. Con un sistema de administración de contenido empresarial, la información no estructurada es gestionada de acuerdo a

reglas de negocio predefinidas, políticas y procedimientos. Es posible establecer relaciones entre piezas de contenido para que el mismo componente pueda ser utilizado en distintos contextos. El sistema provee inteligencia a las colecciones de contenido, creando esquemas de categorización y metadatos que faciliten la búsqueda y recuperación además de la publicación de contenido a través de múltiples interfaces. El sistema asegura el archivamiento y retención a largo plazo para cumplir con requerimientos normativos. Es decir entonces que, los sistemas de Administración de Contenido Empresarial automatizan y administran todo el ciclo de vida de procesamiento del contenido, ver Figura 1. 8.



Figura 1. 8 Sistema ECM

El sistema de ECM permite interactuar a través de una interfaz gráfica de usuario con toda la gama de componentes que facilitan la administración del acervo de contenido como la seguridad, archivadores, metadatos, búsquedas, taxonomía, dinamismo etc.

Fuente Imagen: (EMC, Hoja de Datos h2810.2, 2008)

En el mercado existen varios sistemas de Administración de Contenido Empresarial que se encargan de manera general de cubrir **3 etapas de la Gestión de Contenido Empresarial**:

- Generación y captura de contenido
- Gestionar contenido
- Distribuir y salvaguardar contenido

El elemento núcleo del sistema será el almacén, o repositorio de contenido, donde la información se almacenará de manera estructurada, de tal forma que permita manipular y explotar la información.

A continuación se describen de manera detallada las funciones de las principales etapas de los sistemas de Administración de contenido empresarial.

Generación y captura de contenido:

Esta es la primer etapa de un sistema de gestión de contenidos, en la cual se obtendrán todos los contenidos importantes y serán añadidos al repositorio de contenidos correspondiente con las características que permiten clasificarlos, para esto se analizan las estructuras de datos y de procesos para proceder a mapear sobre las herramientas para que el repositorio tenga la capacidad de albergar las estructuras de datos necesarias para tratar la problemática concreta.

Gestionar contenido:

El repositorio, como ya se ha mencionado, es la base del sistema de gestión de contenido, para ésta segunda etapa ya se tiene el contenido y sus metadatos correspondientes dentro del repositorio, cada uno de los elementos del repositorio puede ser configurado de tal manera que permita:

- Definir los roles asociados al contenido (quién crea, quién aprueba, quién consulta, etc.)
- Modificar el formato del contenido (traducir el contenido, cambiar el formato de un archivo *.doc a, por ejemplo *.pdf, entre otros.)
- Definir el ciclo de vida de un documento; es decir establecer el conjunto de estados por los que atraviesa un documento durante su proceso (por ejemplo, creación, revisión, aprobación, publicación, archivo muerto, y demás.)
- Cumplimiento de ciertos criterios (validar tanto la forma como el contenido de un documento)
- Planificar su publicación, y otros.)

Distribuir y salvaguardar contenido:

Esta etapa se enfoca en publicar el contenido que se encuentra centralizado en el repositorio, el sistema de gestión de contenido permite publicar el contenido por medio de canales muy diversos ya sean documentos impresos, documentos electrónicos compartidos, difusión electrónica por una variedad de foros como web corporativa, portales empresariales o envío directo a otras empresas o clientes. Así mismo el sistema se encarga de salvaguardar el contenido dependiendo de las características que hayan sido configuradas en el ciclo de vida, e incluso puede mejorar la administración de almacenamiento dependiendo de las características que el hardware permita.

En resumen, el sistema de gestión o administración del contenido de una empresa da cierta estructura a la información que por definición es no estructurada; de tal manera que permite gestionar la creación, los procesos, la entrega y el archivado de cualquier contenido según reglas definidas por el negocio, permite establecer relaciones entre fragmentos de contenido, para que éste se utilice en diferentes contextos y bajo distintos formatos. Asimismo, permite añadir inteligencia, creando un esquema de categorización y metadatos que hacen que las búsquedas y recuperación de contenido sean más rápidas y eficaces, automatiza el procesamiento del contenido a lo largo de su ciclo de vida, facilita la publicación de contenidos a través de múltiples canales. La Administración de Contenido Empresarial promueve la integración de los distintos departamentos y de sistemas que anteriormente utilizaban silos de datos, además de permitir la inclusión de la información del contenido no estructurado para que se tome en cuenta para la toma de decisiones de la empresa.

1.5 Documentum - Sistema ECM

Como un ejemplo de los variados sistemas de gestión o administración de contenido empresarial que hay en el mercado y que permiten integrar el modelo holístico de Contenido Empresarial, se encuentra **Documentum de EMC²**, que es una plataforma robusta y flexible que soporta y cumple con las 3 etapas de los sistemas de gestión de contenido (generar, gestionar y distribuir/salvaguardar), para ello cuenta con ciertas características que se lo permiten.

A continuación se muestran las características clave que definen la plataforma de gestión documental de Documentum de EMC²:

Global y distribuida: Para organizaciones que dispongan de clientes y delegaciones por todo el mundo, Documentum maneja a los usuarios y los contenidos independientemente de la ubicación física. La

arquitectura almacena contenido y metadatos en cualquier idioma en repositorios compartidos que conforman un gran repositorio virtual que va más allá de las barreras geográficas e idiomáticas.

Extensible: Documentum puede extenderse para cumplir necesidades de organización únicas incorporando reglas de negocio para objetos específicos. Los plugins personalizables se pueden desarrollar e implementar en áreas clave como la autenticación de los usuarios, manejo de ficheros multimedia o archivado en áreas de almacenamiento.

Abierta: Dado que Documentum está basado en estándares, es muy sencillo de integrar con las infraestructuras IT existentes. Utiliza la arquitectura SOA para exponer sus servicios de gestión de contenido e integrarse con sistemas empresariales. La arquitectura es totalmente fiel al J2EE (para aplicaciones basadas en Web) y soporta completamente el entorno Microsoft .NET y proceso XML. Además, Documentum se integra “out of the box” con aplicaciones y sistemas empresariales como los servicios de directorio utilizando el estándar LDAP.

Escalable: A medida que las necesidades de gestión de contenido aumenten y se hagan más complejas, la solución Documentum manejará de forma eficiente el crecimiento constante de volúmenes de información, cargas de tráfico, aumento de los usuarios y procesos y workflows complejos, todo orientado a lograr el menor coste posible con el mayor rendimiento. Documentum está diseñado para utilizar los recursos de red y dar salida a los problemas de gran escalabilidad a los que se enfrentan las organizaciones hoy en día. La arquitectura de Documentum aprovecha todo el hardware existente para escalarse, aprovechando los sistemas multiprocesador así como los sistemas cache y clusters (escalabilidad horizontal y vertical).

Seguro: Documentum proporciona niveles de seguridad apropiados para que las organizaciones tengan el contenido del repositorio disponible para todos los usuarios y contribuidores. Las Listas de Control de Acceso definen usuarios, grupos y roles que pueden acceder al repositorio y a todos los objetos que lo conforman, además de definir qué tipo de operaciones pueden realizar. La información sensible almacenada en las áreas de almacenamiento se puede encriptar y utilizar comunicaciones seguras de red entre el servidor y los clientes mediante SSL (Secure Sockets Layer). Documentum también soporta firmas electrónicas y ofrece una auditoría extensiva de todas las actividades realizadas en el sistema.

Portátil: Documentum soporta un gran número de bases de datos, sistemas operativos y servidores de aplicaciones y de portales. Los componentes del sistema pueden desplazarse con toda transparencia de una de las configuraciones soportadas a otra.

Documentum admite todas las principales plataformas de bases de datos, sistemas operativos, navegadores, portales, servidores de aplicaciones y estándares de desarrollo como se muestra en la Figura 1. 9. Esta flexibilidad protege la inversión en la infraestructura de TI existente y disminuye el costo general de propiedad.

Soporte de Navegador	Soporte de Servidor de Aplicaciones
Apple Safari	Apache
Microsoft Internet Explorer	BEA
Mozilla Firefox	IBM
Netscape	Oracle
	Sun
Soporte de Base de Datos	Soporte de Plataforma
IBM DB2	AIX
Microsoft SQL Server	HP-UX
Oracle	Linux
Sybase	Microsoft Windows Server
	Solaris

Figura 1. 9 Soporte para entornos de TI

Documentum cuenta con varias opciones de compatibilidad para soportar sus diferentes módulos (navegador, Servidor de aplicaciones, base de datos, y plataforma).

Fuente Imagen: (EMC, Hoja de Datos h2810.2, 2008)

Además de las características anteriores, la plataforma se define para su funcionamiento por:

- Arquitectura lógica
- Arquitectura física

1.5.1 Arquitectura Lógica

La plataforma Documentum provee un ambiente unificado para capturar, almacenar, acceder, organizar, controlar, revisar, liberar y archivar cualquier tipo de información no estructurada dentro de una empresa, esto también soporta las fuentes para manejar el contenido. La plataforma Documentum consiste en cuatro grupos conceptuales como se muestran en la Figura 1. 10:

- *La capa de contenido*, Consiste en el sistema de almacenamiento de archivos y la base de datos que forman el repositorio de Documentum
- *La capa de servicios*, Está compuesta por Documentum Content Server y una variedad de servicios extendidos que otorgan funcionalidad de administración de contenido
- *La capa de interface*, Proporciona una interface de comunicación entre las capa de servicios y la capa del cliente.la comunicación entre la capa de servicios y los clientes que usan esos servicios.
- *La capa de aplicaciones*, Está formada por productos Documentum, aplicaciones integradas y aplicaciones cliente a la medida que se comunican con el Content Server. Cada una de estas capas está compuesta por una serie de componentes los cuales juntos forman un unificada, consistente y extensible arquitectura.

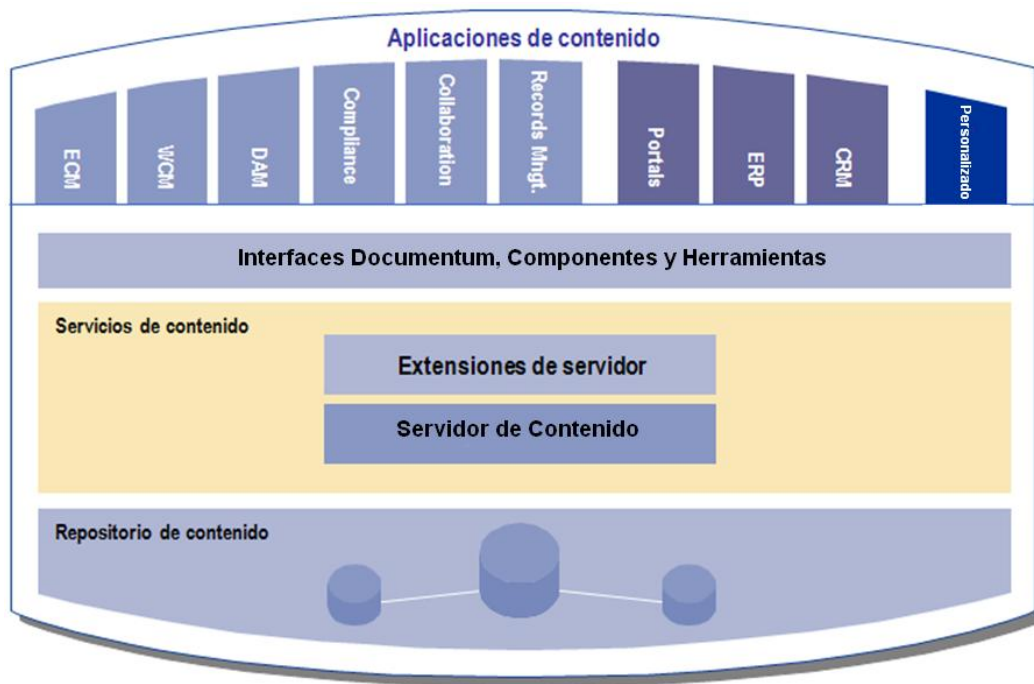


Figura 1. 10 Arquitectura lógica

La Arquitectura lógica Documentum cuenta con 4 capas: La capa de contenido, La capa de servicios, La capa de interface y La capa de aplicaciones; en ése orden debe ser configurada la plataforma.

Fuente Imagen (EMC, Technical Fundamentals of Documentum, 2008)

En la base de la plataforma Documentum se sitúa el repositorio de contenidos empresariales y Documentum Content Server, que gestiona el repositorio e implementa las capacidades de gestión de contenidos esenciales. Content Server pone estas capacidades a disposición de clientes y aplicaciones, a través de la capa de interfaz. Los servicios de gestión de contenidos adicionales se pueden añadir instalando diversas ofertas de servicios de contenidos ampliados.

La capa de contenido:

Los servicios de contenido proveen capacidades fundamentales para poder acceder al contenido almacenado en el repositorio. El repositorio de contenidos de Documentum utiliza un modelo de objeto ampliado para almacenar el contenido y los metadatos asociados, es un repositorio abstracto que consiste en datos almacenados en diferentes fuentes físicas, los documentos están formados por archivos de contenido (el archivo de origen en su formato nativo) y atributos de documento (también conocidos como metadatos o propiedades), como el propietario del documento, la versión y la fecha de creación. Como se puede observar en la Figura 1. 11 el contenido se divide en 2, atributos y contenido propiamente dicho, para facilitar la explotación de las características de cada uno de ellos.



Figura 1. 11 Objeto de contenido

Para su almacenamiento un objeto está formado por el archivo de contenido y los metadatos que lo describen.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

Estos atributos sirven como metadatos que describen el contenido y la relación existente entre este contenido y otros objetos del repositorio. El repositorio utiliza los metadatos para organizar el contenido. Los usuarios pueden usarlo para buscar contenido que les resulte relevante.

Infraestructura del Repositorio

- Almacenamiento de archivos y RDBMS

Típicamente los atributos del contenido son almacenados en un Sistema Manejador de Base de Datos Relacional para que el sistema pueda hacer un query y obtener una respuesta rápidamente, mientras que el contenido activo es guardado en el Almacenador de Archivos; éste puede ser un sistema de archivos host del sistema operativo o un almacenamiento de contenido-direccionado (CAS). A continuación en la Figura 1. 12 se muestra el almacenamiento de objetos

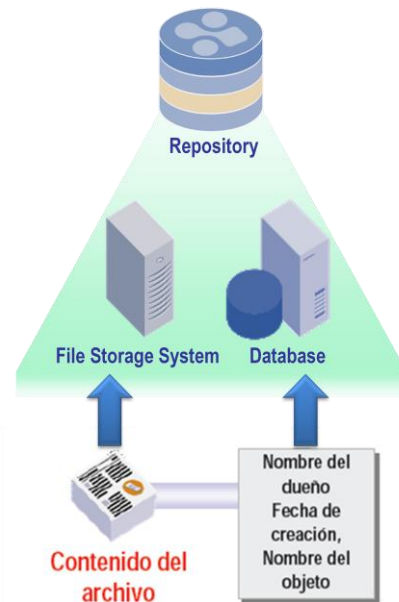


Figura 1. 12 Almacenamiento de Objetos

El repositorio está formado por un Sistema de Almacenamiento de archivos y una Base de datos que albergan a los objetos de contenido; archivo de contenido y metadatos respectivamente.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

- *Full-Text Index*

La plataforma Documentum mantiene un full-text index de todo el contenido activo basado en texto almacenado dentro del repositorio Documentum, permitiendo una rápida búsqueda a través de largas colecciones de información no estructurada, es decir que permite localizar aquel contenido activo por medio de la información dentro del documento.

- *Objetos de Contenido.* El Content Server cumple las medidas de seguridad necesarias para garantizar que sólo los usuarios autorizados pueden acceder a los archivos de contenido, a los metadatos y a los índices. Ningún usuario ni aplicación podrá evadir al servidor para acceder a los datos directamente a través del sistema de archivos o del software de base de datos.

Un objeto es un componente que consiste en datos (archivos de contenido, atributos y relaciones, en el caso de documentos) e instrucciones para las operaciones que se podrán llevar a cabo con esos datos (llamados métodos) como se muestra en la Figura 1. 13. Al igual que el conjunto de atributos, el conjunto de métodos de un objeto se puede configurar y ampliar con las herramientas de desarrollo. Los desarrolladores pueden crear nuevos tipos de objetos que se comporten según sus necesidades empresariales específicas.



Figura 1. 13 Objeto de contenido detallado

Un objeto de contenido además de contenido (texto, imágenes, audio, video, etc.) y metadatos (atributos por defecto y personalizados, formato, etc.) está formado por las instrucciones que se le asignan durante su ciclo de vida (exponer, imprimir, transformar, etc.) y las relaciones que tiene con otros objetos (versión, anotación, etc.).

Fuente Imagen: (EMC, Hoja de Datos h2810.2, 2008)

Las aplicaciones cliente interactúan con los objetos de tipo de documento recurriendo a sus métodos. Desde el punto de vista de la aplicación cliente, los métodos son los mismos independientemente del formato del archivo de contenido subyacente. La encapsulación de las diferencias entre formatos permite que las aplicaciones de gestión de contenido no dependan del formato.

La capa de servicios

Está formada por Documentum Content Server y una variedad de servicios ampliados, ofrece la funcionalidad de gestión de contenidos y sirve como base para todos los demás productos.

Entre los servicios de administración de contenido suministrados por la plataforma Documentum son:

- Servicios de contenido
- Servicios de seguridad
- Servicios ampliados

Servicios de Contenido

Esto incluye servicios de librería, de flujo de trabajo, de ciclo de vida y XML principalmente.

- *Servicios de librería*

Los servicios de librería manejan el contenido de tres formas críticas:

Check in/ Check out (bloqueo): capacidad que asegura que los usuarios que tienen privilegios de edición no editen una versión al mismo tiempo hasta que se hayan actualizado los cambios.

Versiónamiento: capacidad de localizar múltiples versiones de documentos u objetos y provee la habilidad de revertir la prioridad a la versión requerida.

Rendiciones: capacidad de mantener otra alternativa de representar los documentos o u otros objetos de contenido en diferentes formatos, resoluciones u otros lenguajes y mantener la relación entre el objeto original y las rendiciones.

- *Servicios de Flujo de trabajo y BPM*

El proceso empresarial definido consiste en una red conectada de actividades, incluida la información sobre quién lleva a cabo cada actividad y cuándo esto se logra a través de un workflow, que permite guiar un documento entre varios usuarios y tareas automáticas, el workflow es definido por un modelo, la secuencia de pasos que comprende el proceso de negocio, las acciones que deben ocurrir en cada paso. Un flujo de trabajo puede describir un proceso complejo, este puede ser serial con actividades que ocurren una tras otra o en paralelo con actividades que ocurren simultáneamente y se pueden combinar actividades en serie y en paralelo, un estado del workflow es un objeto definido por un conjunto de atributos de contenido adjuntos que viajan con el objeto. En la Figura 1. 14 se muestra un ejemplo de flujo de trabajo.

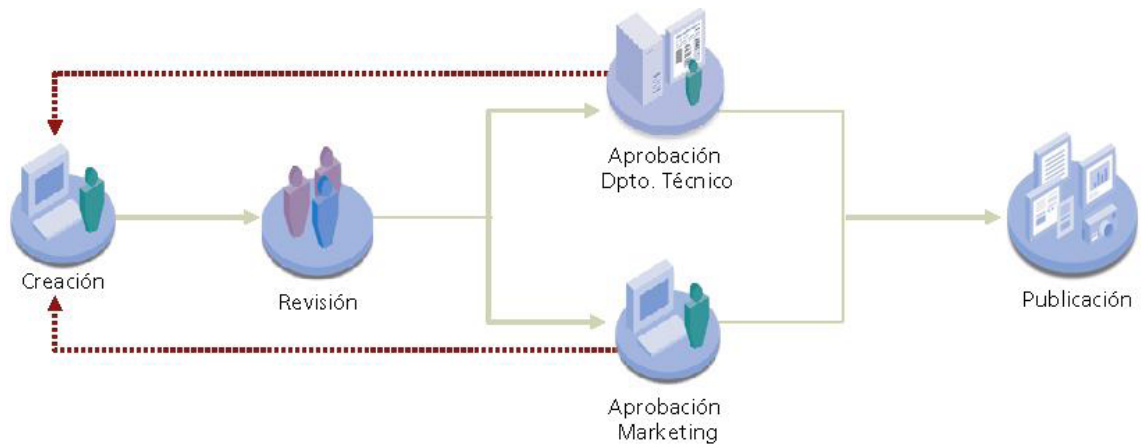


Figura 1. 14 Ejemplo de Flujo de trabajo

El flujo de trabajo y el BPM además de administrar todas las variables del proceso (ciclo de vida, usuarios, roles, almacenamiento, seguridad, objetos de contenido etc.) de tal manera que permita automatizar un proceso de negocio a la medida del cliente.

Fuente Imagen: (EMC, Hoja de Datos h2810.2, 2008)

- ***Servicios de Ciclo de vida***

El ciclo de vida define las reglas de negocio para aplicar los cambios al contenido para mover a través de estados predefinidos (como “borrador”, “en revisión”, “activo” y “obsoleto”). Como un workflow; un estado del ciclo de vida es un objeto está definido por un conjunto de atributos de contenido adjuntos al objeto entonces también viaja con el objeto. Los servicios de ciclo de vida son definidos por un conjunto de políticas o reglas de negocio. Cada contenido debe tener un único ciclo de vida.

Los servicios de Ciclo de vida automatizan las políticas de ciclo de vida del contenido del repositorio. Estos servicios asignan un estado del ciclo de vida al objeto de contenido y entonces manejan la transición de un estado a otro.

- ***Servicios de XML***

XML por sí mismo es una plataforma neutral que separa el contenido del formato, esto tiene muchos beneficios para el manejo del contenido incluyendo habilitar la inteligencia del contenido y la reutilización del mismo. El contenido puede ser requerido por etiquetas predefinidas y valores que permiten una búsqueda con precisión. El contenido puede ser almacenado en una sola fuente y entonces cambiar el propósito y renderizar en múltiples formatos o varios tipos de herramientas de despliegue.

Documentum preserva la estructura jerárquica y liga los componentes XML y los documentos; esto provee la habilidad automática de analizar sintácticamente, validar, transformar, mapear y almacenar dentro de documentos XML. También soporta las aplicaciones XML que directamente se almacena contenido con etiquetas XML (y se maneja el contenido sin ellas) en el repositorio Documentum.

Los servicios XML proveen dos características esenciales para manejar documentos XML en su formato nativo.

Servicios de seguridad

El soporte de la seguridad es provisto por Documentum Content Server (Servidor de Contenido), que incluye: Autenticación, Autorización y Auditado. Primero la plataforma Documentum construye como base la infraestructura de la seguridad de la empresa para autenticar el acceso al repositorio. Siguiendo; la plataforma maneja las listas de control de acceso (ACLs) para autorizar el acceso al contenido almacenado dentro del repositorio. Finalmente, cada actividad puede ser auditada utilizando una flexible herramienta de auditoría que almacena los resultados en el repositorio. La plataforma también puede encriptar todas las comunicaciones entre el Servidor de contenido y otros sistemas como clientes, aplicaciones basadas en web, y cualquier servidor de directorio.

- *Autenticación*

La plataforma Documentum cuenta, inicialmente, con mecanismos de autenticación operando por debajo del sistema o base de datos como se muestra en Figura 1. 15, cada nombre de usuario y password cambia para manejar el acceso al repositorio, la plataforma soporta una autenticación basada en pases para el acceso a nivel aplicación, asegurando que cada aplicación cliente tiene señales válidas para conectar al repositorio y tener acceso al contenido

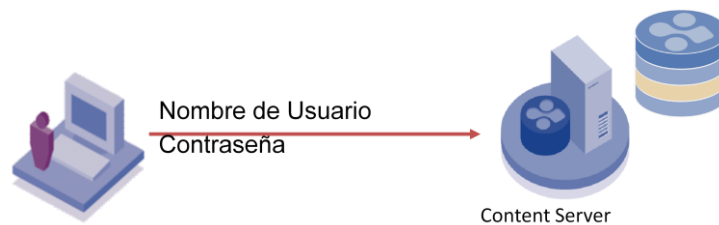


Figura 1. 15 Autenticación

Para acceder al Servidor de contenidos la autenticación de usuarios se realiza a través de una Interfaz Gráfica de Usuario (comúnmente Webtop) y se valida con los registros almacenados en el Servidor de contenido, si es válido el usuario podrá acceder al contenido con los permisos que se le han configurado.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

- *Autorización*

Una vez que el usuario o aplicación autentica su identidad, la persona o programa puede acceder al contenido almacenado basado en privilegios asociados a la identidad. Las reglas de autorización determinan qué contenido puede ser accedido o modificado.

La plataforma Documentum asigna reglas de autorización a través de las Listas de Control de Acceso (ACLs), las cuales son automáticamente aplicadas a los objetos del repositorio cuando éstos son creados. Los ACLs pueden ser modificados manualmente por un usuario o bien automáticamente a través del ciclo de vida, procesos de negocio y otras aplicaciones.

- *Permisos Básicos.*

None: El contenido del repositorio no puede ser visto

Browse: Los atributos del contenido (metadatos) de los objetos de contenido puede ser visto, pero el contenido activo no puede ser abierto o leído.

Read: El contenido activo puede ser abierto y leído, pero no puede ser cambiado.

Relate: Un usuario puede crear relaciones entre el objeto de contenido dado y otros objetos dentro del repositorio (crear comentarios)

Version: Un usuario puede hacer cambios al contenido activo, pero no puede sobrescribir una versión existente.

Write: Un usuario puede hacer cambios al contenido de un objeto o a los metadatos asociados a él y guardar los cambios sin tener que generar una nueva versión. Este nivel de control de acceso es usualmente confidencial para el dueño del objeto.

Delete: Un usuario puede borrar el contenido de un objeto.

Este set de permisos es acumulativo cada nivel automáticamente garantiza el acceso a los derechos de los niveles anteriores a excepción de “Delete”, ya que solo tendrá el permiso de borrar el contenido de un objeto sin tener permiso de “versión”, “relate”, “read” y “browse”.

- *Permisos extendidos*

Change location: Un usuario puede cambiar la locación de un contenido activo de un folder a otro.

Change permission: Un usuario además del dueño del contenido puede cambiar los permisos estándar del contenido activo.

Change owner: Un usuario, además del dueño del contenido, puede cambiar el dueño del contenido activo.

Execute Procedure: Un usuario puede ejecutar un procedimiento externo en un contenido activo como crear una rendición.

Change estate: Un usuario puede cambiar el estado del ciclo de vida del contenido activo.

- *Auditado*

Cada operación realizada por el repositorio de Documentum puede ser grabada en un registro auditable. La pista de la auditoría puede ser configurada en Documentum Administrator (donde también puede ser vista) y es asegurada en el repositorio por una fuerte encriptación.

Cada lista de registro auditable, ambas, los nuevos y previos valores asociados con un evento (como el tiempo, nombre de usuario, cuando un documento es puesto en “check out” del repositorio), habilitan la rápida determinación de lo que ha cambiado. Los usuarios finales y administradores pueden también ver el historial de los documentos y otros objetos almacenados en el repositorio para determinar cómo y cuándo la información cambia.

Servicios Ampliados

Los productos de servicios ampliados de Documentum aumentan las capacidades de Content Server más allá de sus servicios de gestión empresarial esenciales. Estos productos pueden ampliar cualquier aspecto del marco de trabajo de gestión de contenido.

La capa de interfaz

Los clientes y las aplicaciones utilizan la capa de Interfaz para comunicarse con el Content Server e interactuar con el repositorio de contenido. La capa de interfaz como se muestra en Figura 1.16 está formada por Documentum Foundation Classes (DFC) y por cierto número de interfaces estándar construidas sobre DFC. Conjuntamente, estos productos componen la Interfaz de Programación de Aplicaciones Documentum (DAPI).

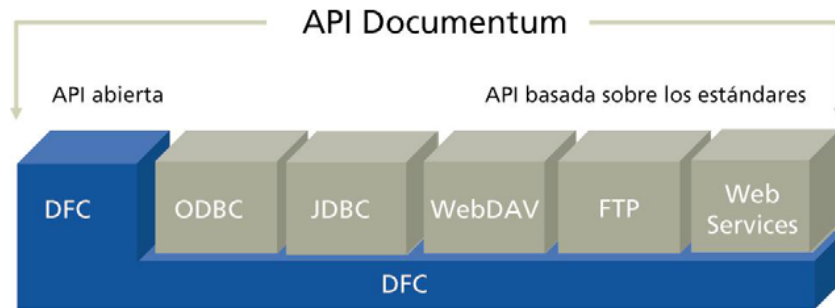


Figura 1. 16 La capa de interfaz

La capa de interfaz de la plataforma Documentum está basada en DFC) que son las clases programadas en Java que le dan la funcionalidad a la plataforma. A partir sus estándares se puede utilizar ya sea ODBC, JDBC, WebDAV, FTP, Web Services o bien hacer configuraciones de autor.

Fuente Imagen: (EMC, Hoja de Datos h2810.2, 2008)

Documentum Foundation Classes, es la API más compleja que integra todas las funcionalidades de Documentum. DFC proporciona un marco de trabajo orientado a objetos para acceder a las funcionalidades del Servidor de Contenido. DFC expone el modelo de objeto Documentum como una biblioteca cliente orientada a objeto para ser utilizada por las aplicaciones de gestión de contenidos. Esta API puede ser utilizada desde un host de aplicaciones, incluyendo aplicaciones desarrolladas en Java, Visual Basic, C y C++. Ver Figura 1. 17.

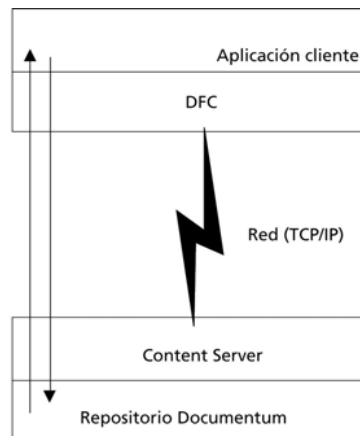


Figura 1. 17 Conectividad con aplicación cliente

El usuario podrá acceder al contenido almacenado en el Repositorio Documentum desde la Aplicación cliente, de manera transparente al usuario hace uso de la conectividad; llamando las clases Documentum DFC, la información viaja a través de la Red (TCP/IP) hasta el Servidor de contenido y a través de los permisos administrados por éste podrá manipular el contenido haciendo uso de la misma conectividad en ambos sentidos.

Fuente Imagen: (EMC, Hoja de Datos h2810.2, 2008)

Las aplicaciones acceden al Servidor de Contenido a través de DFC, utilizando un objeto cliente. Al crear un objeto cliente se cargan las bibliotecas compartidas necesarias ver Figura 1. 17 Conectividad con aplicación cliente .Las interfaces estándar comprenden:

- **ODBC y OLEDB** — Muchas herramientas para la elaboración de informes, como Crystal Reports y Microsoft Access, se benefician de los protocolos de acceso a datos de Microsoft para comunicarse con el repositorio Documentum.

- *JDBC*—Muchas aplicaciones de servidor utilizan el protocolo estándar de acceso a datos Java para acceder al contenido del repositorio Documentum.
- *WebDAV*—Documentum WebDAV Services ofrece un servidor WebDAV que permite la utilización de aplicaciones compatibles con WebDAV, como Adobe Photoshop y Documentum Desktop para Macintosh, que emplean este protocolo para comunicarse con el repositorio Documentum.
- *FTP*—Documentum FTP Services es un servidor FTP para el repositorio Documentum que permite que herramientas como Macromedia Dreamweaver se integren en el repositorio gracias al protocolo de transferencia de archivos estándar de Internet.
- *Web Services*. La lógica del negocio, puede exponerse como Web Services, otorgando una manera de llamar funciones en otras computadoras a través de una intranet o Internet. Por ejemplo, un sistema CRM puede comunicarse con Documentum a través de Web Services.

La capa de aplicaciones

La capa de aplicación cliente se compone de aplicaciones básicas para que los usuarios accedan al repositorio Documentum. Incluye interfaces para usuario final, herramientas de administración e integraciones con las herramientas de autor (personalizadas) y las aplicaciones empresariales más comunes.

Las aplicaciones de la capa cliente se forman de componentes reutilizables que se basan en los servicios de gestión de contenidos disponibles en las capas inferiores de la arquitectura. Una aplicación cliente de Documentum es una colección de componentes que operan juntos para llevar a cabo un objetivo empresarial. Los mismos componentes pueden figurar en otras aplicaciones clientes o como portlets en el portal de una empresa. En la Figura 1. 18 se muestra la estructura de la capa de aplicación cliente.

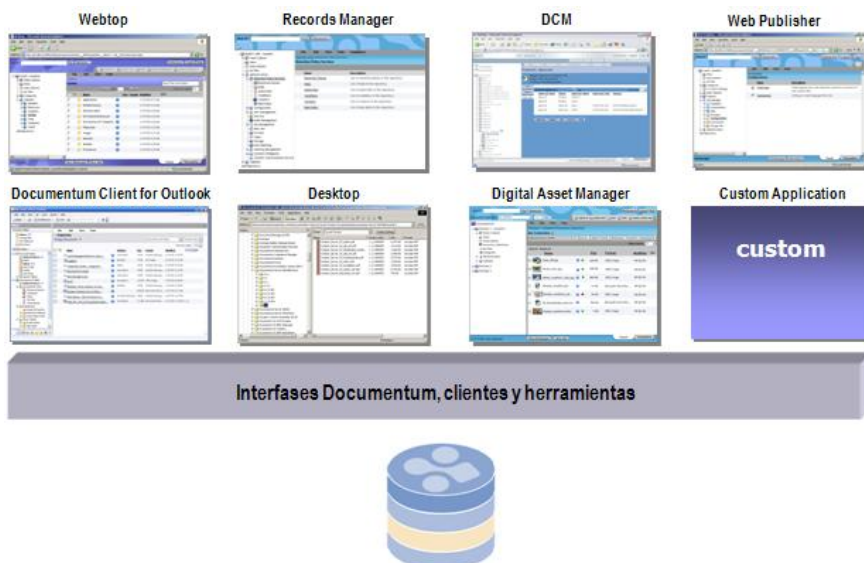


Figura 1. 18 Capa de aplicación cliente

La capa de aplicación reside sobre la capa de interfaz. El objetivo de esta capa es permitir al usuario la interacción directa con el contenido almacenado en el repositorio; existen varias aplicaciones cliente que la plataforma Documentum soporta (Webtop, Records Manager, DCM; Desktop, Aplicaciones personalizadas, etc.) y que permiten administrar el sistema y el contenido.

Fuente Imagen: (EMC, Hoja de Datos h2810.2, 2008)

Aplicaciones basadas en el entorno de Microsoft Windows

Documentum Desktop es una aplicación Microsoft Windows que provee a los usuarios finales acceso a uno o más repositorios Documentum y presenta todas las funcionalidades de gestión de documentos empresariales. Se integra en Windows Explorer, presentando contenido del repositorio en una interfaz paralela al acceso de Explorer a contenidos de sistemas de archivos como se muestra en la Figura 1. 19.

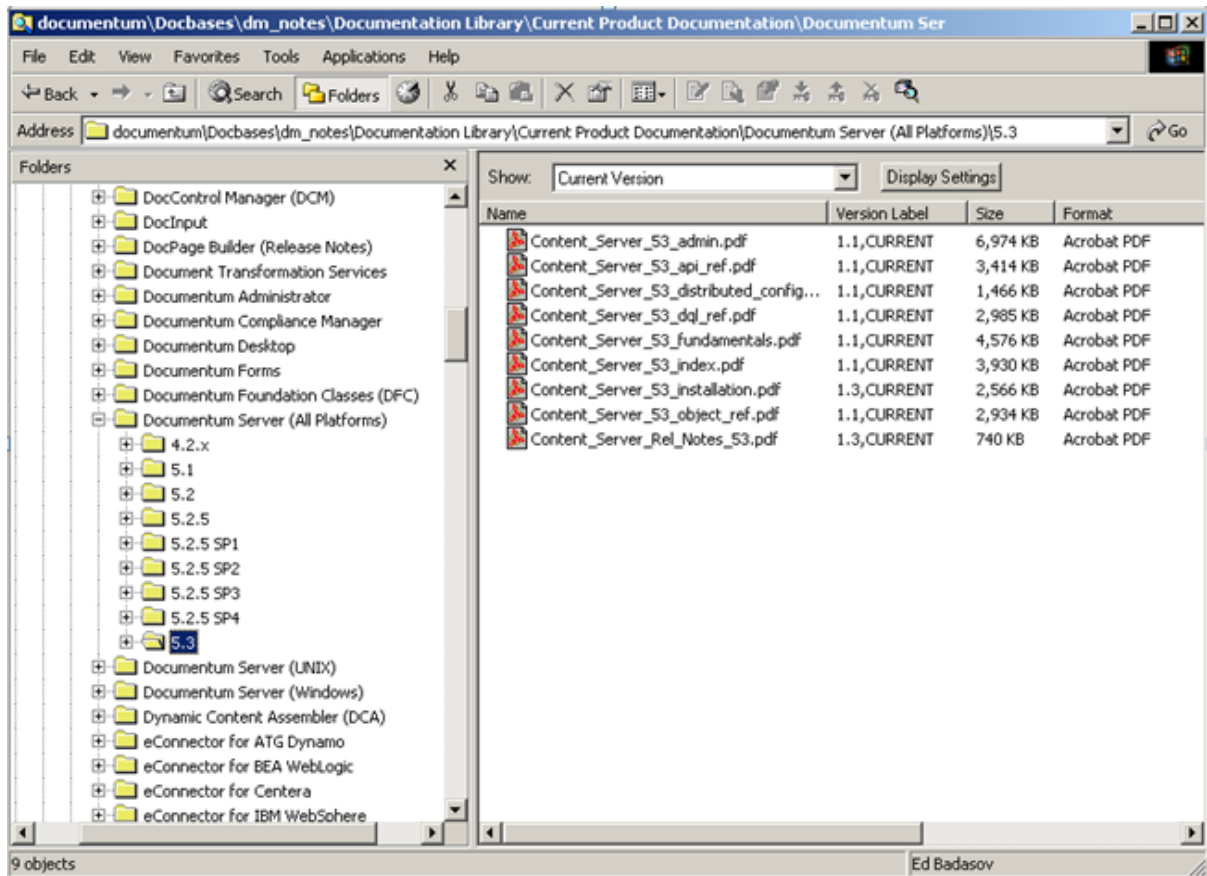


Figura 1. 19 Documentum Desktop

Entre las aplicaciones cliente que soporta la plataforma Documentum se encuentra Documentum Desktop que es una versión muy similar en vista y funcionalidad al escritorio de un ordenador.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

Documentum Desktop también provee integración en aplicaciones comunes de otros fabricantes como Microsoft Office, para que puedan utilizar servicios de gestión de contenidos directamente desde la interfaz de usuario de la aplicación. Entre otras, Documentum Desktop permite integrar:

- Windows Explorer
- Microsoft Office
- Microsoft Outlook
- Acrobat Business Tools y Exchange

Los desarrolladores pueden personalizar los componentes de Documentum Desktop o utilizarlos en aplicaciones personalizadas. La capa de aplicaciones permite integrar herramientas de

generación de contenidos en Documentum tanto en plataformas Macintosh como Windows; incluye un componente de servidor que permite una unidad de acceso a la red a través del sistema operativo del cliente. Este servidor permite que el sistema operativo trate el repositorio Documentum como si fuese un sistema de archivos sin poner en peligro su seguridad. Desde la perspectiva de las herramientas de autor (personalizadas), estas leen y escriben archivos de un disco.

Aplicaciones Documentum basadas en Internet

Documentum Webtop es una aplicación basada en Internet que procura a los usuarios finales acceso a uno o más repositorios Documentum. Provee un entorno de usuario similar al de Documentum Desktop, empleando un navegador en lugar de Microsoft Windows Explorer. La siguiente Figura 1. 20 muestra cómo luce Documentum Webtop.

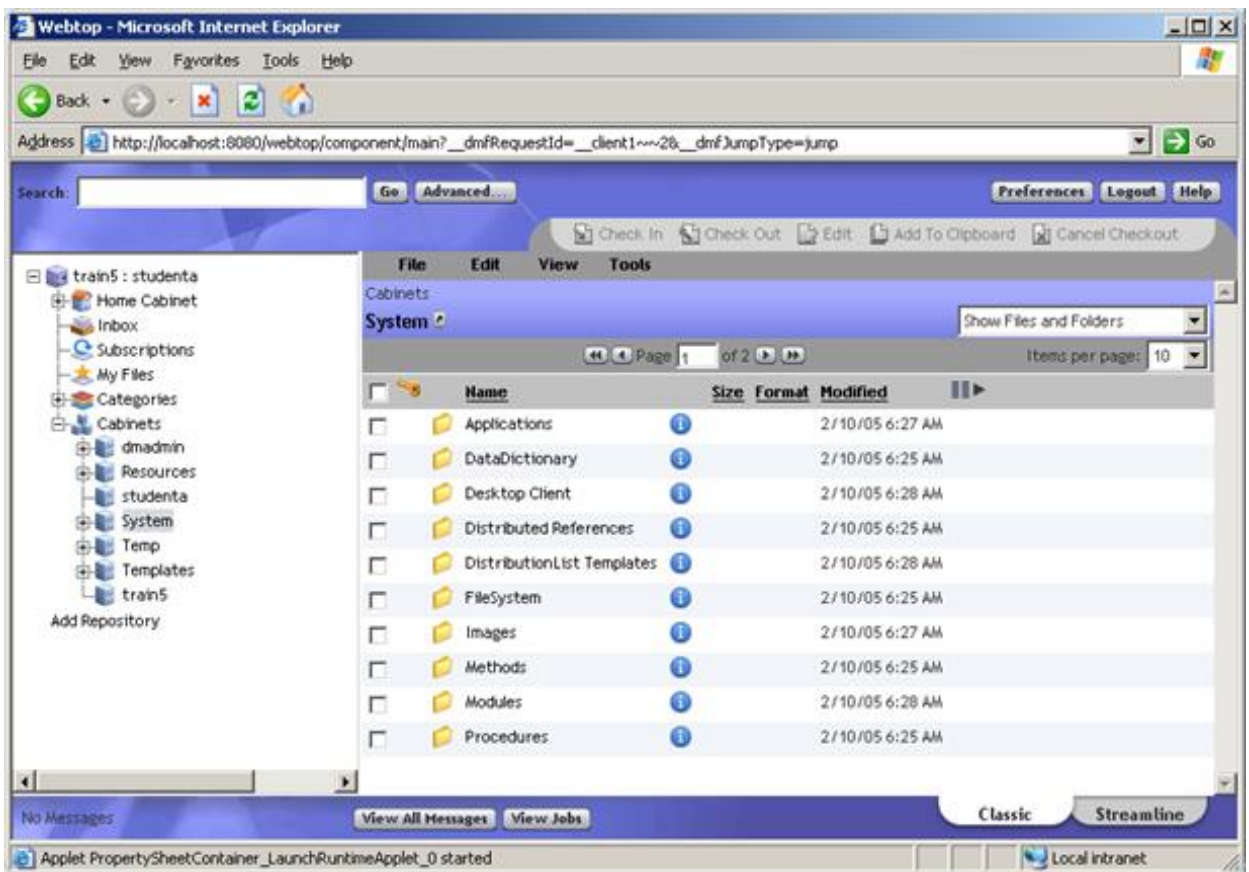


Figura 1. 20 Documentum Webtop

Entre las aplicaciones cliente que soporta la plataforma Documentum se encuentra Documentum Webtop que en vista es similar a un buscador de archivos de un ordenador y en la funcionalidad existe una bandeja de entrada para las tareas asignadas en un proceso, la gestión de contenido, usuarios, grupos, roles, búsquedas etc.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

Webtop incluye un servicio de caracterización que permite a los usuarios personalizar las interfaces de usuario de sus aplicaciones. Este servicio gestiona el aspecto de la interfaz de usuario mediante temas que incorporan imágenes, iconos y hojas de estilo en cascada. Los controles de la interfaz de usuario, los encabezados de las etiquetas, pueden ser configurados

empleando una hoja de estilos en cascada e indicando la ubicación de los archivos de imágenes utilizados para dar forma a los controles. Los usuarios pueden seleccionar uno de los temas disponibles. En entornos multilingües, los usuarios pueden seleccionar el idioma elegido para la interfaz de usuario cuando se conectan.

Webtop ha sido creado utilizando Documentum Web Development Kit (WDK). De la misma forma en que los desarrolladores pueden incorporar componentes de Documentum Desktop a una aplicación de Windows (pre instalada o personalizada), pueden utilizar componentes WDK en aplicaciones Web personalizadas.

Documentum Administrator, la aplicación que administra centralmente Content Server, es una aplicación basada en WDK que permite a los clientes crear, gestionar y publicar webs empleando las funcionalidades de gestión de contenidos de Documentum. Es la interfaz que permite al usuario administrar el Content Server.

Integración del Servidor de Contenido en aplicaciones empresariales

Diversos productos que permiten interactuar con el sistema de gestión de contenidos de Documentum desde otras aplicaciones empresariales.

- Documentum Content Services for SAP
- Documentum Content Services for Siebel eBusiness Application
- Documentum Content Services for Lotus Notes Mail
- Documentum Content Services for Portales

El Servidor de Contenido enlaza contenido de los repositorios Documentum con objetos ubicados en la aplicación empresarial, brindando a los usuarios acceso a funcionalidades clave de Documentum mediante la conocida interfaz de la aplicación empresarial, ahorrando tiempo y costes asociados a búsquedas, archivado y almacenamiento de documentos.

Entre las aplicaciones más recurrentes con las que se puede integrar Documentum se encuentra SAP. Los usuarios pueden revisar procedimientos operativos estándar, notas con información sobre seguridad, dibujos de ingeniería y especificaciones en un formulario maestro de material desde SAP. Documentum permite emplear procesos basados en reglas para vincular contenidos empresariales, desencadenar flujos de trabajo y notificar a los usuarios finales que se han producido modificaciones relevantes de contenido. Automatizar el proceso de enlace permite vincular fácilmente grandes volúmenes de contenido escaneado o introducido a distancia, como facturas por cobrar, a registros ERP nuevos o existentes, garantizando una captura de datos sin errores y la validación de los procesos.

1.5.2 Arquitectura Física

La arquitectura lógica que se desee configurar en la plataforma Documentum no puede ser llevada a cabo hasta que todos los componentes se encuentren instalados y la conectividad entre ellos sea favorable, es por esta razón que a continuación se muestran las conexiones básicas del Servidor de Contenido y otros componentes.

Establecer una conexión

La arquitectura física está dividida en tres capas:

- Primera capa: Interfaz cliente
- Segunda capa: Lógica del negocio
- Tercera capa: Almacenamiento de datos

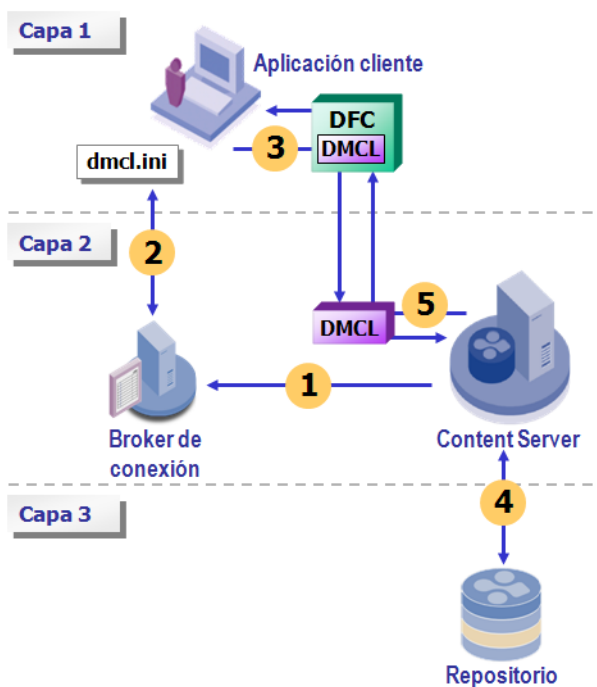


Figura 1. 21 Comunicación con aplicaciones cliente - servidor

La comunicación cliente – servidor:

- 1.- El Broker de conexión reconoce al Servidor de contenido.
- 2.- La aplicación cliente se conecta al Broker de conexión.
- 3.- El usuario a través de la aplicación cliente establece la conexión con el Servidor de Contenidos a través de una biblioteca de interfaces DMCL y un framework orientado a objetos DFC.
- 4.- El Servidor de contenido procesa la información solicitada por el usuario, accede al repositorio y obtiene la información.
- 5.- El Servidor de contenido envía la información solicitada vía DFC y DMCL.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

De acuerdo con la Figura 1. 21 la comunicación entre las 3 capas tiene un flujo determinado, mismo que se explicará a continuación.

- 1.- Se reconoce un Servidor de Contenido a través del Broker de conexión; éste almacena en su memoria información sobre el Content Server.

- Nombre del servidor
- Id del servidor de procesos
- Nombre de la máquina host
- Estado del servidor
- Fecha y hora en la que el Broker de conexión recibió información la última vez del servidor
- Cuando el Broker de conexión espera escuchar nuevamente al servidor
- Nombre del repositorio (no necesariamente uno)
- Id del repositorio

Los Content Servers pueden registrarse con múltiples Brokers de conexión. Cada cliente especifica un Broker de conexión en el archivo dmcl.ini

2.- El cliente se conecta al bróker de conexión para obtener la información de conexión al Content Server

3.- El cliente establece la conexión y envía comandos a la API del Content Server a través de DFC. Una API (Application Programming Interface) es un conjunto de instrucciones que se ejecutan en el servidor para realizar una función específica. En orden de ejecutar APIs en el servidor, los clientes usan DMCL. DMCL soporta el protocolo RPC (Remote Procedure Call) para permitir a los clientes entrar al Content Server sin entender detalles de la red.

La API del Content Server, es referida como un DMCL, es como una biblioteca con interfaces para funcionalidad en todos los servidores

El DFC es un framework orientado a objetos utilizado para ingresar a las capacidades del Content Server. Cada cliente corre una aplicación cliente como una copia del DFC corriendo en una Java Virtual Machine (JVM)

4.- El content server procesa los comandos del cliente y obtiene la información solicitada del repositorio

5.- El Content Server envía la información de regreso a la API cliente vía DFC

Por lo general esta comunicación se ve modificada cuando se utiliza un modelo de componentes que encapsulan la funcionalidad del Content Server; el framework WDK, ver Figura 1. 22.

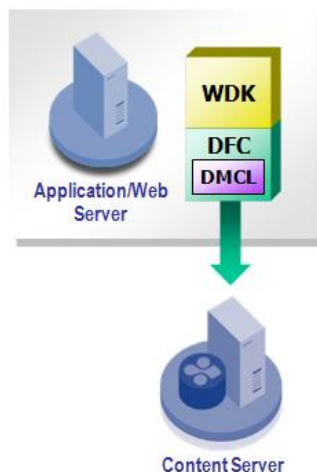


Figura 1. 22 Aplicaciones WDK

La aplicación WDK permite desarrollar aplicaciones web personalizadas encapsulando el funcionamiento del Servidor de contenido, la comunicación cliente-servidor se ve ligeramente modificada al utilizar WDK, puesto el uso del framework orientado a objetos DFC, y la librería de interfaces DMCL, ahora serán accedidos a través del framework WDK, que se integrará y configurará a la medida a través de un Servidor de aplicaciones.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

El framework WDK permite un rápido desarrollo de aplicaciones web que ingresen a los repositorios de Documentum. Incluidos con WDK hay muchos componentes pre-construidos que son fáciles de configurar e integrarse a la aplicación. Las aplicaciones web a la medida

pueden basarse en WDK o extenderse a otras aplicaciones WDK clientes como Webtop. El modelo de desarrollo WDK se basa en las tecnologías J2EE y XML. La capa de presentación usa etiquetas HTML y librerías para el control de la aplicación en las páginas web

Páginas Web WDK

Una página jsp en WDK es compuesta de una plantilla fija de HTML y contenido dinámico que es presentado por la tecnología JSP. La lógica de la aplicación está separada de la página JSP en clases de bibliotecas. Los eventos de la interfaz del usuario son trasladados en llamadas a métodos de objetos del lado del servidor.

Las páginas JSP se compilan en servlets (clases de java) por un contenedor JSP o un compilador de terceros. Estos servlets se ejecutan en el servidor de aplicación Java y ejecutan tareas de servidor o generan contenido dinámico que se despliega en el visor cliente.

De tal forma que la comunicación como se aprecia en la Figura 1. 23 queda de la siguiente manera:

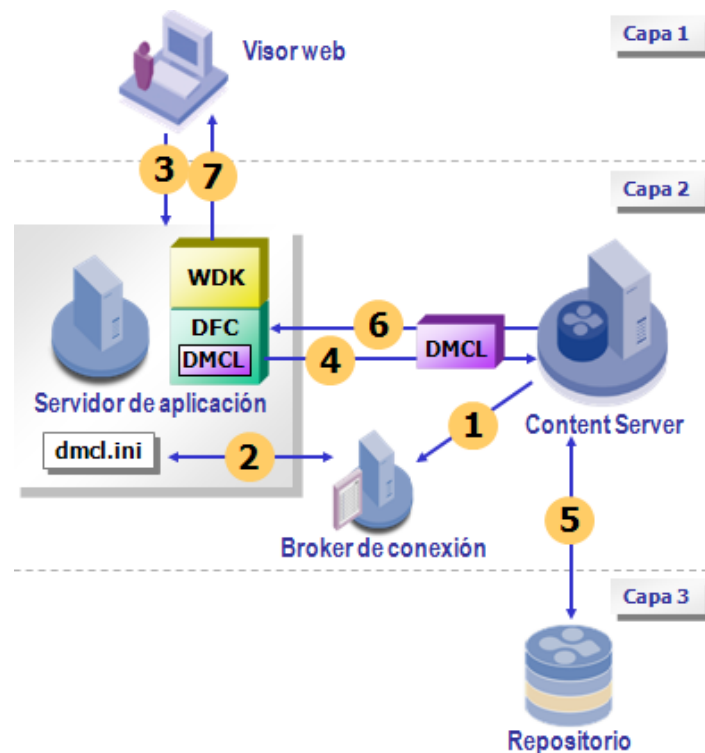


Figura 1. 23 Comunicación con aplicaciones web

La comunicación cliente – servidor por medio de una aplicación web.

- 1.- El **Broker de conexión** reconoce al **Servidor de contenido**.
- 2.- El **Servidor de aplicaciones** se conecta al **Broker de conexión**.
- 3.- El usuario se conecta al **Servidor de aplicaciones** a través de una URL que se resuelve con JSP.
- 4.- El **Servidor de aplicaciones** a través de métodos WDK hace referencia al framework orientado a objetos DCF y a la biblioteca de interfaces DMCL para establecer la conexión con el **Servidor de Contenidos**.
- 5.- El **Servidor de contenido** procesa la información solicitada por el usuario, accede al **repositorio** y obtiene la información.
- 6.- El **Servidor de contenido** envía la información solicitada vía DFC y DMCL al **servlet** del **Servidor de aplicaciones**.
- 7.- El **servlet** envía la información al **cliente** a través del **visor web**.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

El flujo de la información mostrado en la Figura 1. 23 se describe como sigue:

- 1.- Content server se registra por sí mismo con el Broker de conexión
- 2.- El servidor de aplicación se conecta con el Broker de conexión para obtener la información de conexión del Content Server
- 3.- El cliente usa un URL para el servidor de aplicación El servidor de aplicación resuelve la solicitud a la página JSP correcta El motor JSP crea dinámicamente un servlet desde la página JSP
- 4.- El servlet usa métodos WDK para llamar a las clases DFC o a Objetos del negocio que usan comandos de la API del Content Server
- 5.- El Content Server procesa las APIs y recupera la información solicitada del repositorio
- 6.- Content Server envía la información de regreso al servlet
- 7.- El servlet envía el contenido al visor web.

Configuraciones del repositorio:

Ahora bien ya que se ha definido la manera en que se lleva a cabo una conexión, es importante explicar las diferentes configuraciones en que los repositorios, Servidores de contenido y clientes se relacionan como se visualiza en la Figura 1. 24.

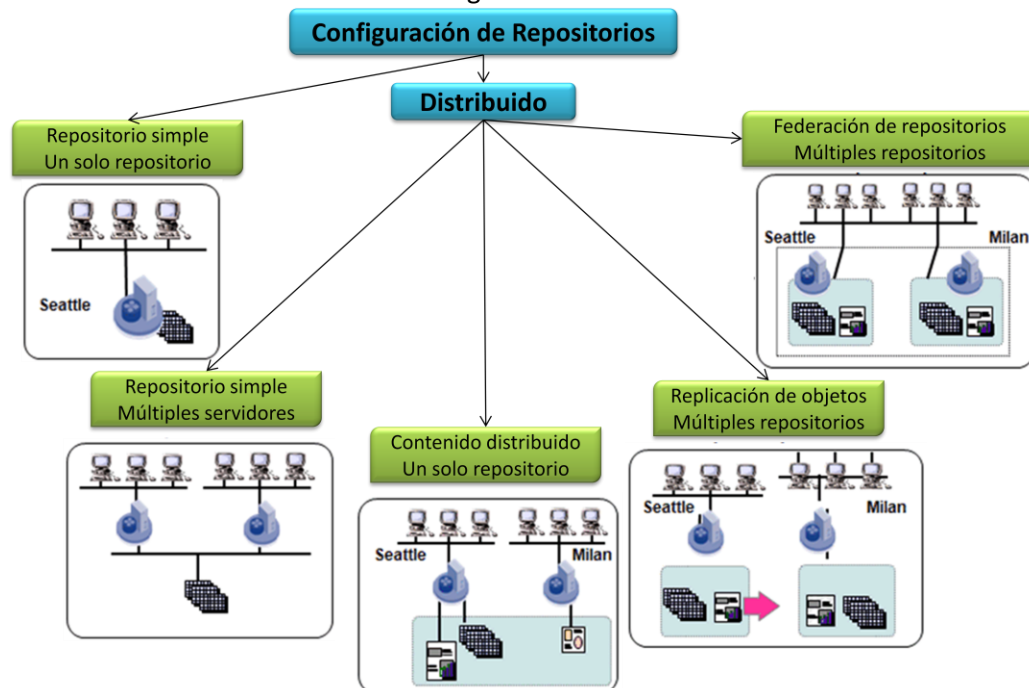


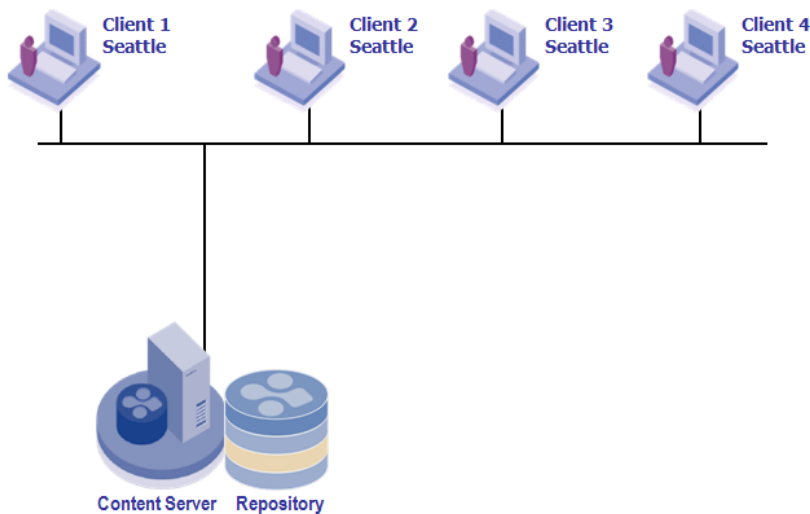
Figura 1. 24 Configuraciones del repositorio

El contenido es almacenado por el RDBMS y este a su vez es administrado por el Servidor de Contenido; para satisfacer las necesidades de demanda de información por cliente existen dos configuraciones para el repositorio;

- 1.-Repositorio simple con un solo repositorio.
- 2.-Repositorio distribuido el cual a su vez tiene 4 configuraciones;
 - a) Repositorio Simple
 - b) Contenido distribuido
 - c) Replicación de objetos
 - d) Federación de repositorios.

Fuente Imagen: (Solís, 2010)

Repositorio simple; Un servidor, un repositorio: Se utiliza un solo Servidor de contenido y un solo repositorio; hay que tener en cuenta que varios clientes accediendo a un mismo repositorio disminuyen el tiempo de respuesta. Ver Figura 1. 25.



La configuración básica recomendada para una demanda de usuarios baja, dentro de una red local es; un Servidor de contenido y un repositorio. Todos los objetos (contenido y metadatos) son consultados de manera local.

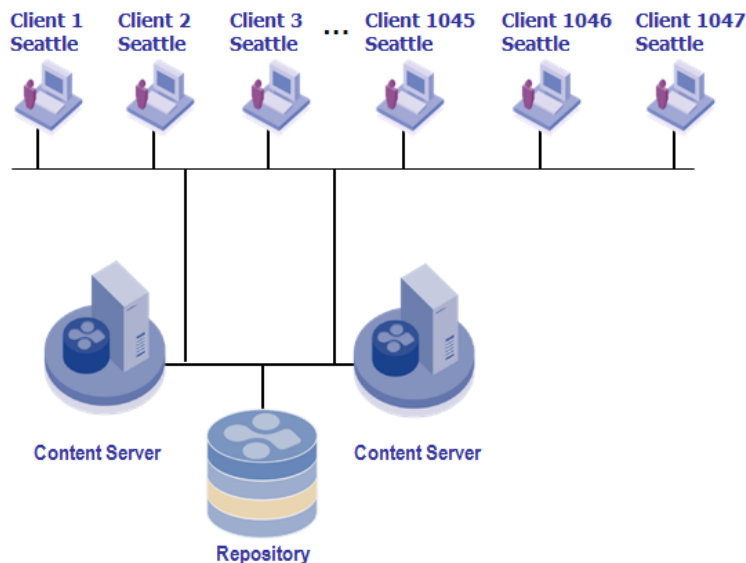
Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

Figura 1. 25 Un servidor, un repositorio

Repositorio simple, múltiples servidores:

Se utilizan múltiples servidores de Contenido y un solo repositorio como puede observarse en la Figura 1. 26; esto se hace para poder distribuir la carga de sesiones de usuario ya que se coloca dentro de los servidores de contenido localmente donde se necesita. Esta configuración puede mejorar el performance cuando muchos usuarios simultáneos acceden a documentos grandes:

- Los documentos grandes tienen la misma cantidad de metadatos que los documentos pequeños
- La carga en el servidor de base de datos no es impactada por usuarios accediendo a documentos grandes; la carga en el Content Server, quien entrega los archivos, si es impactada por esta actividad.



Cuando la demanda de información es mayor dentro de una zona geográfica limitada, es necesario instalar más Servidores de contenido; quienes accederán al mismo repositorio dividiendo la carga de sesiones entre ellos.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

Figura 1. 26 Múltiples servidores

Contenido distribuido: Como se aprecia en la Figura 1. 27, en esta configuración el contenido de un sistema de almacenamiento de archivos es copiado a uno o más sistemas de almacenamiento de archivos. Por lo que muchos Servidores de Contenido tienen acceso local al mismo documento y hay mejoras de rendimiento:

- Cuando los usuarios acceden documentos grandes de ubicaciones remotas
- La administración de redes se reduce, ya que no se requieren protocolos para compartir almacenamiento

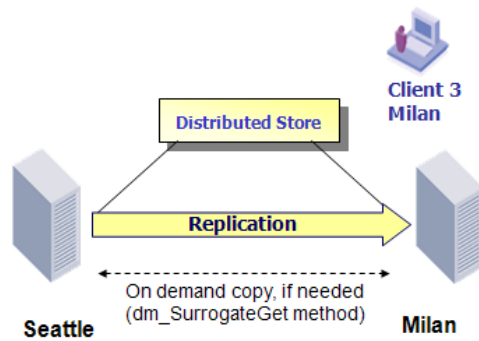


Figura 1. 27 Replicación de contenido

Cuando los límites geográficos se superan, la entrega de contenido se vuelve lenta, así que es necesario replicar el contenido de un objeto y colocarlo de manera tal que el acceso en distintas localidades geográficas sea ágil.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

Replicación de objetos: Es una configuración en la cual los objetos (metadatos y contenido) son copiados de un repositorio a otro, por lo tanto los clientes utilizan un solo repositorio para datos y contenido permitiendo a los usuarios en diferentes ubicaciones tener acceso rápido al mismo documento. Las réplicas pueden ser modificadas cuando se editan, el objeto fuente también se actualiza para reflejar los mismos cambios ya sea en propiedades como en contenido. Ver Figura 1. 28,

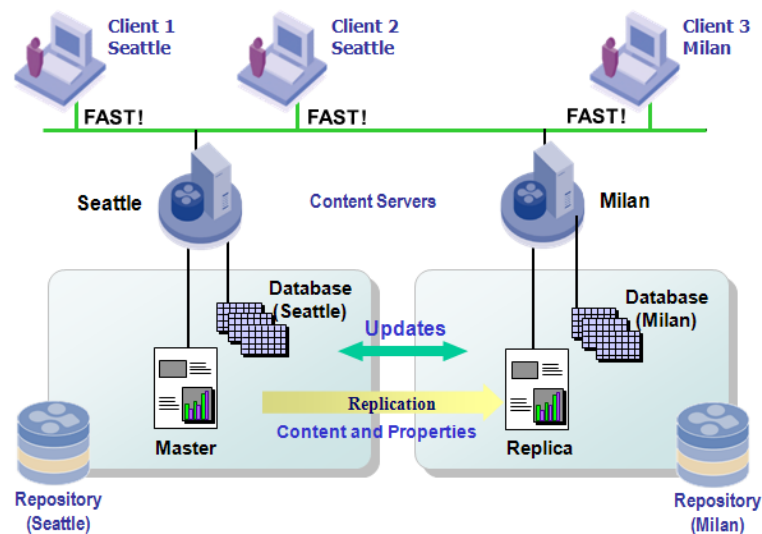


Figura 1. 28 Replicación de objetos

Cuando los límites geográficos son superados, la demanda de un mismo contenido es recurrente por ambos sitios geográficos es necesario que se replique todo el objeto, no solo el contenido sino también los metadatos, así que cada Servidor de contenido tendrá su correspondiente repositorio pero los objetos que almacenan los repositorios es exactamente la misma y debe automatizarse la actualización de un objeto en ambos repositorios.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)

Federación de repositorios: Es una configuración en la que un grupo de repositorios se encuentran unidos para manejar usuarios, grupos y conjunto de permisos externos de manera global; la configuración se conforma por un repositorio principal y repositorios miembros. El repositorio principal describe la relación entre un repositorio y los otros repositorios dentro de una federación, controla los usuarios globales, grupos y conjuntos de permisos, un repositorio puede pertenecer a una y sólo una federación. Sólo puede haber un repositorio principal, pero tantos repositorios miembros como se necesiten. Ver Figura 1. 29.

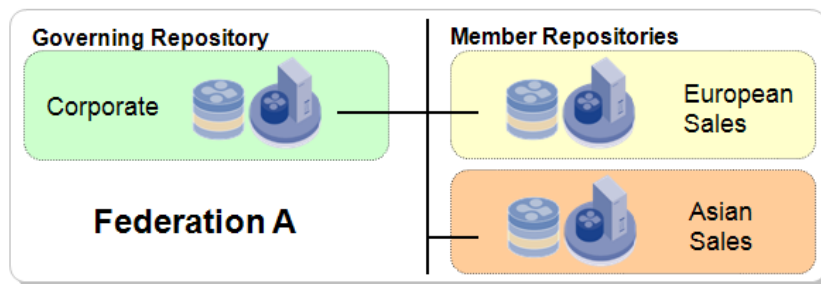


Figura 1. 29 Federación de repositorios

Cuando los límites geográficos se superan y es necesario tener el control de todos los accesos, por parte de una sola localidad, se opta por instalar varios servidores de contenido con sus respectivos repositorios sin tener que replicar los objetos necesariamente (Repositorios miembro), pero siendo administrados los usuarios, roles, grupos y permisos por un Servidor de contenido con su propio repositorio, (Repositorio principal). Ambos repositorios, los miembros y el principal forman una Federación de repositorios.

Fuente Imagen: (EMC, Technical Fundamentals of Documentum, 2008)