

Tema 6. Sistemas de Archivo Distribuido (DFS)

○ Tiene las mismas funciones que el sistema de ficheros de un SO convencional, pero más complejo

○ Usuarios y dispositivos están dispersos por la red

○ Permite

- Compartir información remotamente
- movilidad de los usuarios
- Disponibilidad

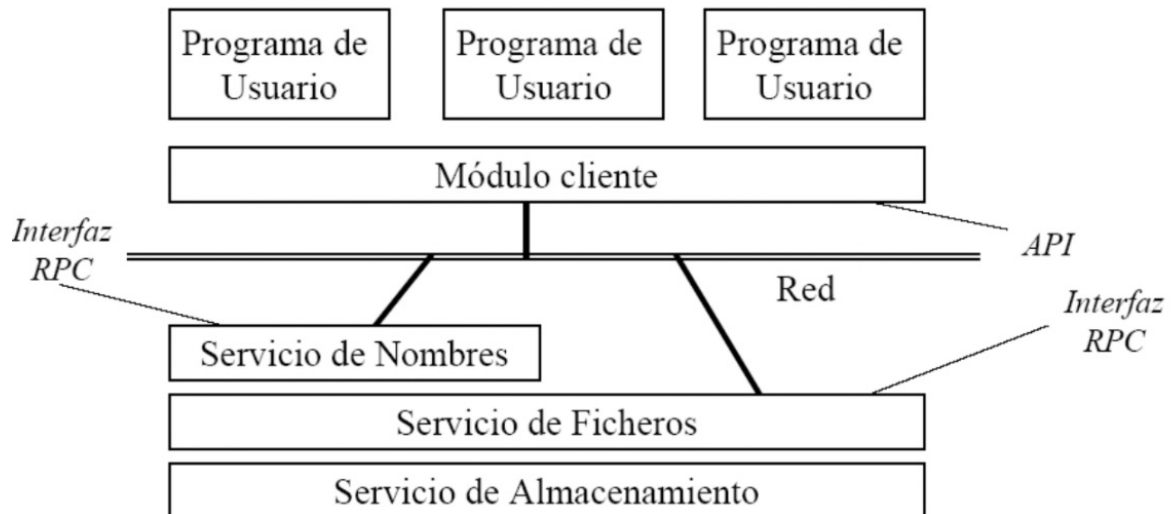
○ Estructura

servicio de almacenamiento.

servicio de ficheros.

servicio de nombres o de directorio.

módulo cliente (API)



→ Objetivos del diseño

○ Transparencia

- acceso
- movilidad
- escalas
- ubicación
- prestaciones

○ Movilidad de usuarios

○ Rendimiento

○ Alta disponibilidad y tolerancia a fallos

○ Concurrencia

○ Fiabilidad e integridad de la información

○ Seguridad

○ Heterogeneidad (HW's y SO's)

* Hardware y sistemas operativos

○ Modelos básicos

Servicio Remoto

- operaciones en los servidores
- Problemas de eficiencia

Cache de datos

- ficheros locales.
- aumento del rendimiento.
- Problemas de consistencia.



combinados → NFS, DropBox, OneDrive...

○ Según servidores implementación de los DFS:

- Caché: copia asociada al cliente a nivel de bloque
- Replicación de ficheros: copia asociada a servidor a nivel de fichero completo.

○ Servidores:

Sin Estado (Stateless)

- Toleran fallos
- No requiere llamadas para abrir o cerrar ficheros
- No se desperdicia memoria en tablas
- No hay límite de ficheros en uso
- No hay problemas si que un cliente

Con Estado (stateful)

- Mensajes de solicitud de servicio más cortos
- mejor rendimiento
- hay operaciones de lectura anticipada
- se puede bloquear ficheros

Localización de caché

- 1.- Disco del servidor (No aconsejable, muy costoso)
- 2.- Memoria del servidor

Ventajas:

- Transparente para los clientes
- Los SO tradicionales suelen usarlo (NFS)

Desventajas:

- Coste de las transferencias por red

- 3.- Disco del cliente (puede no tener disco)

Ventajas:

- fiabilidad
- gran tamaño de caché

Desventajas:

- hay que acceder al disco

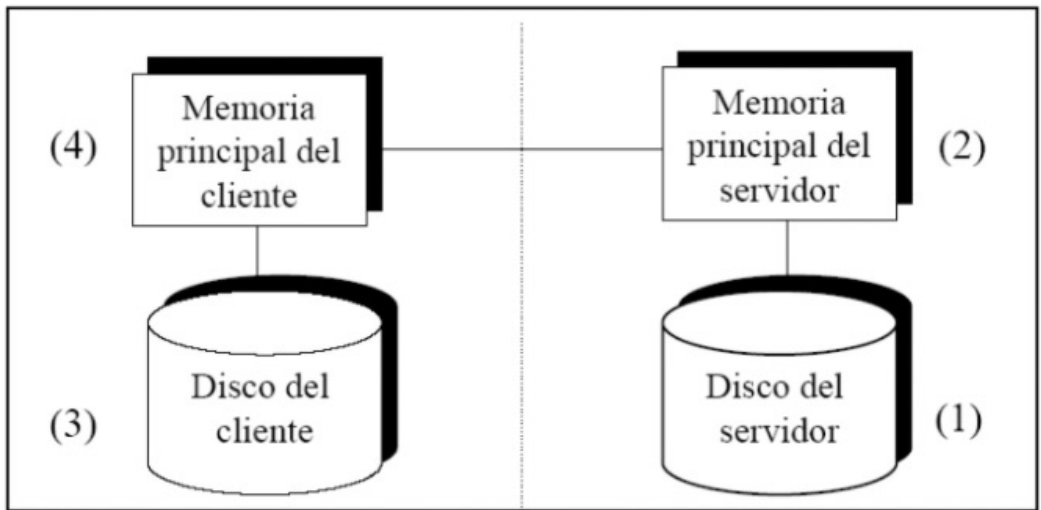
- 4.- Memoria del cliente

Ventajas:

- No hay costes de acceso a disco ni de transmisión de red
- Si se accede máximo rendimiento

Desventajas:

- inconsistencia de información en la caché



Tratamiento de la caché

- En clientes se pueden producir inconsistencias
- Aspectos a considerar:
 - Cuando propagar las modificaciones de la caché al fichero
 - Como verificar los datos de la caché
- Esquemas básicos para propagar las modificaciones.
 - Escritura inmediata (Write-through)
 - Proponer escrituras (Delayed-Write o Writeback)

Validación

- Política de propagación de modificaciones
 - Especifica cuando se actualiza la copia principal de un fichero cuando una de las copias de la caché es modificada
 - No establece cuándo actualizar las cachés

○ El contenido de la caché se vuelve inválido cuando otro cliente modifica la copia principal.

- Se ha de comprobar que la caché de un cliente es consistente con la copia principal
- Si no se ha de invalidar la caché y actualizar datos

○ Estrategias

- iniciada por los clientes
- iniciada por los servidores

Ficheros Replicados: aquel del que existen varias copias en diferentes servidores

○ Ventajas:

- más disponibilidad
- más fiabilidad
- mejor tiempo de respuesta
- menos tráfico de red
- mejor rendimiento (throughput)
- beneficia la escalabilidad
- usable en operación desconectada

○ Cuestiones a considerar:

- Replicación transparente a los usuarios
- Actualizar copias en caso de modificar una réplica

Transparencia de replicación:

○ Aspectos a considerar

- Denominación de las replicas

- Como distinguir replicas con mismo ID

- Un servidor de nombres debería corresponder al id la réplica más conveniente

- Control de la replicación

- Explícita

- Implícita

+ Replicación perezosa

+ Comunicación a grupos

Actualización de réplicas (Protocolos):

○ Replicación de Lectura

- en ficheros inmutables

○ Escribir en todos - Leer en cualquiera:

- Las escrituras en costosas

- Problemas si un servidor cae

○ Basado en la disponibilidad de copias

- problemas de inconsistencia si se parte la red

○ Copia primaria

- se lee de cualquiera, se escribe en una

- si cae servidor primario \rightarrow problemas

○ Buscado en Quorum

- los clientes necesitan quorum N_r servidores para leer y N_w para escribir

- $N_r + N_w > N$ (N : nº de réplicas)

OneDrive for Business Synchronization

How does OneDrive for Business synchronize files?

The business of synchronization

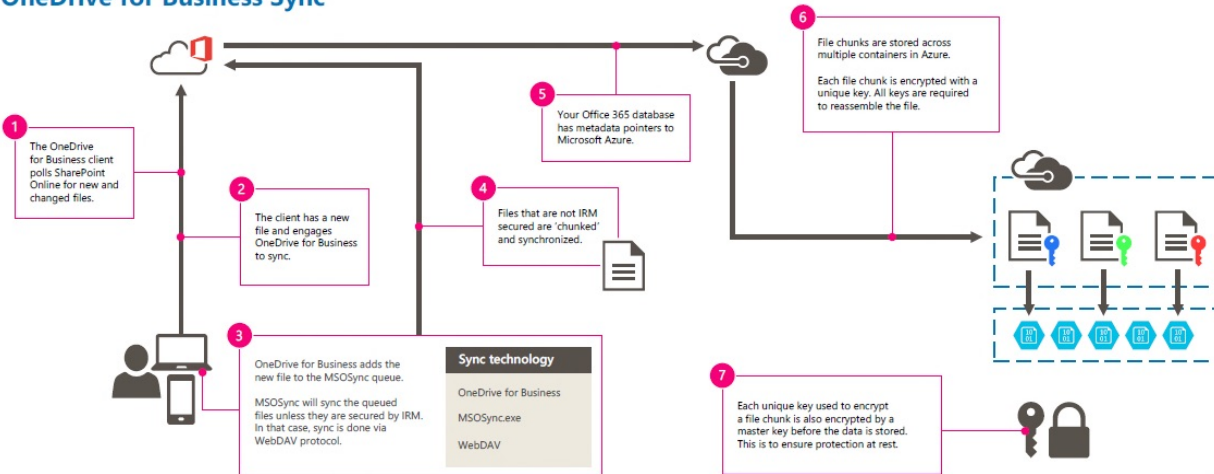
OneDrive for business is constantly checking for new and changed files to synchronize. When a change is detected, such as a new file on a client machine, client synchronization determines the technology to use to sync. Web services are used to bridge the gap between the client computer and Office 365.

Web Services used in synchronization

- **Cellstorage.svc** - used for file synchronization.
- **Sites.asmx** - returns site collection information.
- **Webs.asmx** - used to work with sites and webs.
- **Lists.asmx** - used to work with lists and list data.
- **Version.asmx** - used to work with file versions.



OneDrive for Business Sync



© 2014 Microsoft Corporation

Microsoft

