# Arquitectura de Oracle DataBase.

**Oracle Database** es, en pocas palabras, un servidor que maneja un montón de datos en un entorno multiusuario, permitiendo accesos simultáneos, bloqueando accesos no autorizados y facilitando la recuperación en caso de fallos.

**Base de datos vs. Instancia:** El servidor Oracle se compone de una "base de datos" y al menos una "instancia".

* **Base de datos:** Son archivos en disco que guardan los datos del usuario. Pueden existir sin una instancia y, en la versión 21c, se refiere a archivos de bases de datos multitenant (CDB), bases de datos conectables (PDB) o contenedores de aplicación.
* **Instancia:** Es un conjunto de estructuras de memoria que administra esos archivos. Incluye un área de memoria compartida (SGA) y varios procesos en segundo plano, y puede existir sin archivos de base de datos.

Arquitectura multitenant

**Arquitectura multitenant:** Oracle permite tener un **CDB (Container Database)** que contiene varias bases de datos, llamadas **PDBs (Pluggable Databases)**. Esto significa que una base de datos Oracle puede contener otras bases de datos dentro de ella.

**Componentes clave:**

* **CDB:** Es un conjunto de archivos (control, redo log, y de datos) administrado por la instancia de base de datos. Puede contener múltiples PDBs y contenedores de aplicaciones.
* **PDB:** Es una base de datos “portátil” con su propio set de archivos de datos, que se conecta o desconecta del CDB. Para moverla o archivarla, debe desconectarse del CDB.

**Oracle Database 21c:** Solo admite arquitectura multitenant, es decir, que cada base de datos debe estar en un CDB. Antes, se admitían bases de datos sin contenedor (no-CDB).

**Contenedores de aplicaciones:** Son CDBs específicos de aplicaciones que almacenan datos y metadatos comunes a varios PDBs de esa aplicación. Cada **application root** (base raíz de aplicación) puede incluir definiciones compartidas como tablas o usuarios. En SaaS, se pueden crear PDBs para cada cliente, que comparten metadatos pero mantienen datos independientes.

Instancia de Oracle

* **Instancia de Oracle:** Es el conjunto de **procesos** y **estructuras de memoria** que necesita el servidor Oracle para funcionar. Cada instancia incluye dos estructuras clave de memoria:
  + **SGA (Server Global Area):** Memoria compartida entre todos los procesos del servidor. Contiene datos globales de la instancia, como cachés y "pools" (fondos de memoria reservados para operaciones específicas).
  + **PGA (Program Global Area):** Memoria específica para cada proceso servidor. Si hay varias conexiones, cada una tiene su propia PGA. Incluye datos que aceleran las operaciones SQL, como variables de sesión y espacio para consultas SQL en curso.
* **Componentes internos en la PGA:**
  + **Espacio de pila:** Siempre se guarda en la PGA, incluso en modo de servidor compartido.
  + **UGA (User Global Area):** Área de usuario que puede almacenarse en la SGA en modo compartido. Incluye:
    - **Área de trabajo SQL:** Acelera el procesamiento de SQL con áreas específicas (ordenación, hash, bitmaps).
    - **Memoria de sesión:** Guarda variables y control de sesión.
    - **Área privada SQL:** Almacena el estado de las instrucciones SQL actuales y usa un cursor para gestionar los datos devueltos en consultas SQL largas.
* **Modo de servidor compartido:** Algunos datos de la UGA pasan a la SGA, dejando solo el espacio de pila en la PGA.

### **Explicación de componentes adicionales:**

* **Pool Compartido (Shared Pool):** Memoria en la SGA que guarda instrucciones SQL ya ejecutadas y otros datos compartidos. Ayuda a reducir la carga al reutilizar instrucciones ya procesadas.
* **Caché de Búferes de Datos (Database Buffer Cache):** También en la SGA, es donde se almacenan temporalmente los datos que se leen o modifican para agilizar el acceso a datos usados frecuentemente.
* **Búfer de Redo Log:** Guarda cambios recientes a los datos de la base antes de escribirlos en disco, permitiendo recuperación en caso de falla.
* **Pool Grande (Large Pool):** Área de memoria adicional en la SGA que maneja operaciones más pesadas, como tareas de recuperación y copias de seguridad, para evitar que el Pool Compartido se sature.
* **Pool de Java:** Sólo se usa para los programas que utilizan Java.
* **Pool de Streams:** Lo usa sólo el componente Oracle Streams (utilizado para bases de datos distribuidas) y sirve para almacenar en búferes, datos manejados por dicho componente.

### **Tipos de Procesos en Oracle:**

1. **Procesos de Usuario:** Conectan a usuarios con la base de datos y lanzan sus instrucciones.
2. **Procesos de Base de Datos:**
   * **Proceso Servidor:** Atiende a cada usuario y ejecuta sus consultas SQL o PL/SQL.
   * **Procesos en Segundo Plano:** Se lanzan con la instancia y realizan tareas específicas.
3. **Procesos de Aplicación:** Servicios que se ejecutan de forma automática, como el listener de red.

### **Tareas de los Procesos Servidores:**

* Analizan el código SQL o PL/SQL.
* Definen el plan de ejecución.
* Leen los datos requeridos desde la SGA o el disco.
* Devuelven los resultados o indican los errores.

### **Procesos en Segundo Plano:**

1. **DBWn (Database Writer Process):**
   * Escribe bloques modificados (dirty buffers) de la caché de búferes de datos en los archivos de datos del disco en puntos de control (checkpoints). Puede haber hasta 20 DBWn según la configuración del servidor.
   * Tipos de checkpoints:
     + **Incremental:** Graba solo algunos bloques sucios para liberar espacio.
     + **Total:** Graba todos los bloques sucios y ocurre con menor frecuencia para evitar corrupción en caso de errores.
2. **LGWR (Log Writer Process):**
   * Graba en los archivos Redo Log cada vez que se confirma una transacción o cuando se llena un tercio del búfer de Redo Log. También actúa antes de DBWn y cada tres segundos para asegurar la recuperación de datos en caso de fallo.
3. **CKPT (Checkpoint Process):**
   * Marca la llegada de un checkpoint actualizando la posición en los archivos de control y comunica a DBW la necesidad de grabar los búferes sucios en el disco.
4. **SMON (System Monitor):**
   * Recupera la instancia en su arranque, limpia segmentos temporales y restaura datos no grabados de Redo Log en archivos de datos.
5. **PMON (Process Monitor):**
   * Supervisa los fallos de procesos de usuario, reinicia servidores y libera recursos cuando fallan los procesos de usuario.
6. **RECO:** Solo para bases de datos distribuidas. Resuelve fallos en transacciones distribuidas.
7. **MMON y MMNL:**
   * **MMON:** Realiza tareas relacionadas con el área de estadísticas de rendimiento (AWR).
   * **MMNL:** Guarda estadísticas en disco desde el histórico de sesiones activas (ASH).
8. **ARCn (Archiver Processes):** Archiva los Redo Logs históricos para recuperar información en modo ARCHIVELOG.
9. **CJQ0 y Jnnn (Job Queue Processes):**
   * **CJQ0:** Programa tareas y lanza los procesos Jnnn que ejecutan trabajos planificados.
10. **FBDA (FlashBack Data Archiver Process):**

* Graba información para la funcionalidad Flashback, permitiendo restaurar la base de datos a un estado anterior en caso de necesidad.

Estos procesos en segundo plano ayudan a optimizar y asegurar la integridad y recuperación de la base de datos en Oracle, manejando almacenamiento, rendimiento y registro de transacciones en segundo plano.

Estructuras de almacenamiento de bases de datos

### **Estructuras de Almacenamiento Físico en Oracle:**

Oracle almacena los datos y su información de respaldo en varios tipos de archivos físicos, cada uno con un propósito específico para asegurar la consistencia, recuperación y gestión de la base de datos.

1. **Archivos Redo Log**:
   * Los archivos redo log almacenan un registro secuencial de todos los cambios realizados en la base de datos. Son esenciales para la recuperación de datos, ya que registran las modificaciones desde el último checkpoint.
   * Los redo logs permiten que cualquier transacción se pueda revertir en caso de fallo del sistema, utilizando los cambios guardados en ellos.
   * La base de datos usa múltiples archivos redo log en un grupo de redo log, alternando entre ellos en un ciclo (log switch) para mantener los cambios siempre respaldados y disponibles para recuperación.
2. **Histórico de Archivos Redo Log (Archive Redo Log)**:
   * Estos son copias de los archivos redo log que se generan cuando la base de datos está en modo ARCHIVELOG.
   * Los archive redo logs son cruciales para la recuperación a un punto en el tiempo y el mantenimiento de datos históricos, ya que mantienen un historial de transacciones pasadas.
   * Los archivos redo log se archivan automáticamente al llenarse para asegurar que los datos históricos queden registrados sin interrumpir el ciclo de los redo logs actuales.
3. **Otros Tipos de Archivos**:
   * **Archivos de Datos (Data Files):** Almacenan las tablas, índices y demás objetos de usuario de la base de datos.
   * **Archivos de Control (Control Files):** Contienen información de control de la base de datos, como la estructura de la base de datos y el estado de los redo logs. Son esenciales para la recuperación.
   * **Archivos Temporales (Temporary Files):** Utilizados para operaciones que requieren almacenamiento temporal, como ordenar y agrupar, sin afectar los datos permanentes.
   * **Archivos de Desbordamiento (Undo Files):** Guardan datos anteriores de una transacción en curso para permitir la reversión de cambios si la transacción no se completa.
4. **Posibilidades de Gestión Física de los Archivos de Datos**:
   * Oracle ofrece varias opciones para la organización física de los archivos de datos en el sistema:
     + **Ubicación de Archivos en Directorios Específicos:** Se pueden almacenar en directorios seleccionados según los requerimientos de espacio y redundancia.
     + **Asociación con Tablespaces:** Los archivos de datos están agrupados en tablespaces, lo que permite organizar los datos por áreas lógicas y asignarles espacios físicos.
     + **Distribución en Múltiples Discos:** Para mejorar el rendimiento y la disponibilidad, los archivos de datos pueden ubicarse en diferentes discos físicos o RAID, ayudando a evitar la pérdida de datos ante una falla.
     + **Replicación y Seguridad (Multiplexing):** Oracle permite la replicación de ciertos archivos como los archivos de control y los redo logs, lo cual asegura su disponibilidad en caso de fallo.

Estas estructuras físicas de almacenamiento y las opciones de gestión de Oracle brindan una arquitectura sólida para la seguridad, respaldo y recuperación de los datos, manteniendo la base de datos en operación continua y segura frente a errores.

Estructuras de almacenamiento lógicas

Las **estructuras de almacenamiento lógico** en Oracle Database permiten una gestión eficiente del espacio en disco, dividiéndose en varios niveles:

1. **Tablespaces**:
   * Son estructuras lógicas que agrupan segmentos y conectan los niveles lógico y físico de la base de datos.
   * Oracle crea tablespaces por defecto, como USERS (para objetos de usuario), SYSTEM (para objetos del sistema), y SYSAUX (para componentes adicionales).
   * Un tablespace puede contener varios archivos de datos y se divide en segmentos, extensiones y bloques.
2. **Segmentos**:
   * Cada tablespace contiene segmentos, que corresponden a objetos como tablas, índices o tipos LOB (CLOB y BLOB).
   * Existen segmentos de usuario, de anulación (undo) para revertir cambios, y temporales para datos intermedios en consultas complejas.
   * Un segmento puede estar distribuido en varios archivos de datos, facilitando el almacenamiento de objetos grandes.
3. **Extensiones**:
   * Los segmentos se subdividen en extensiones, que aseguran bloques contiguos de disco y reducen la fragmentación.
   * Cuando un segmento se llena, se añade una extensión para ampliar su tamaño y reservar más espacio contiguo.
4. **Bloques de Datos**:
   * Son la unidad más pequeña de almacenamiento de datos en Oracle y se configuran según el tamaño de bloque del sistema operativo.
   * Los bloques almacenan datos en filas y contienen una cabecera que identifica la tabla y filas almacenadas, con espacio reservado para crecimiento interno del bloque.

Estas estructuras lógicas optimizan el almacenamiento y administración de los datos, asegurando un acceso eficiente y organizado en Oracle Database.