



Arquitectura de Oracle



Objetivos de la Lección

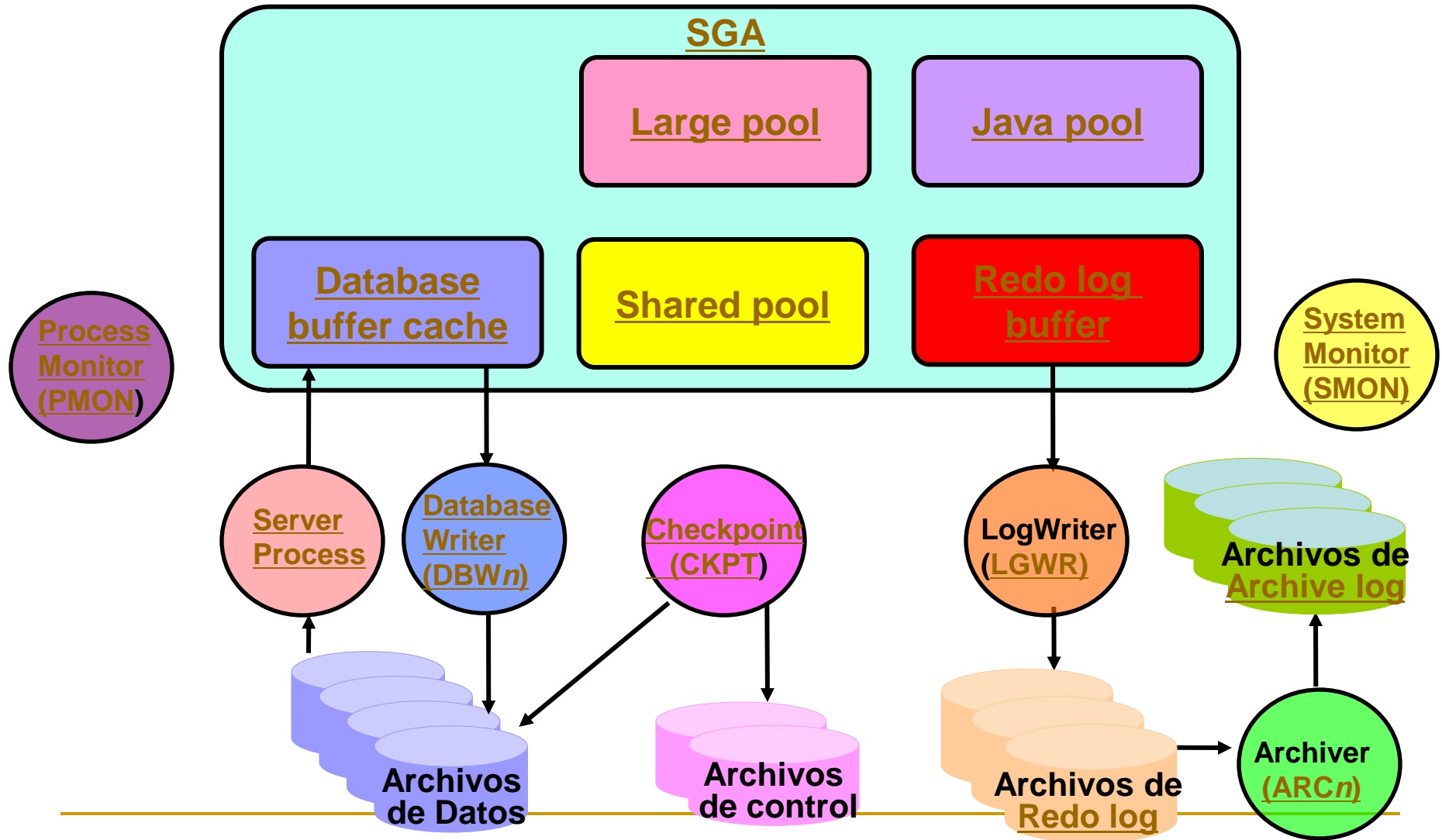
- Definición del Servidor de BD
 - Funcionamiento de la Instancia
 - Estructuras de Memoria de la Base de Datos
 - Estructura de los Procesos de la Base de Datos
 - Establecimiento de conexión y creación de una sesión
 - Estructuras de Almacenamiento de la Base de Datos
-

Servidor de BD de Oracle

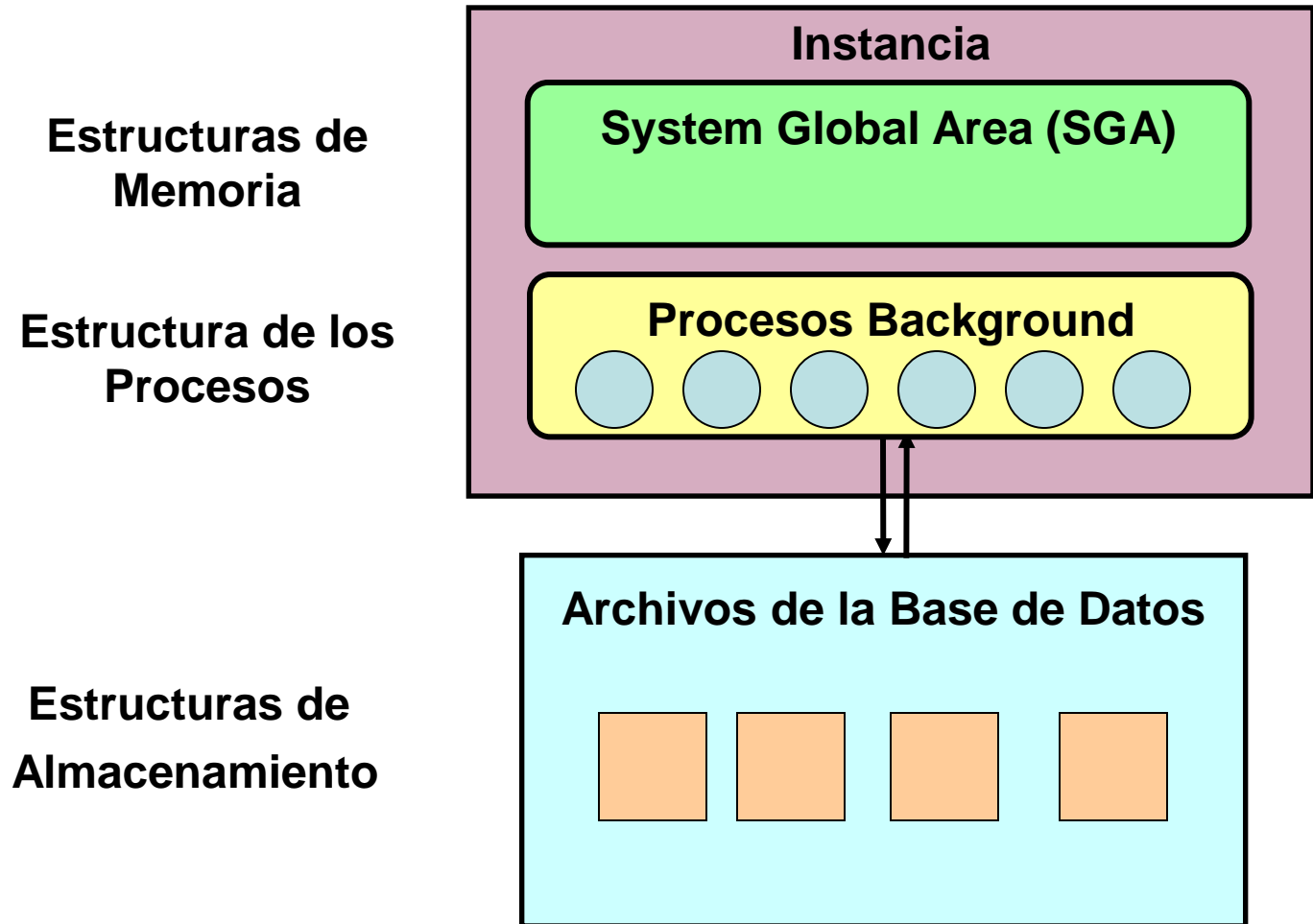
■ Consiste de:

- Una instancia de la BD: combinacion del System Global Area (SGA) y procesos *background* de Oracle. Multiples instancias pueden ejecutarse concurrentemente en el mismo computador, c/u accedendo su propia BD física.
- Una base de datos: cuando se inicia es asociada a una instancia (montar la base de datos). La BD puede ser abierta para ser usada por multiples usuarios autorizados

Funcionamiento de la instancia



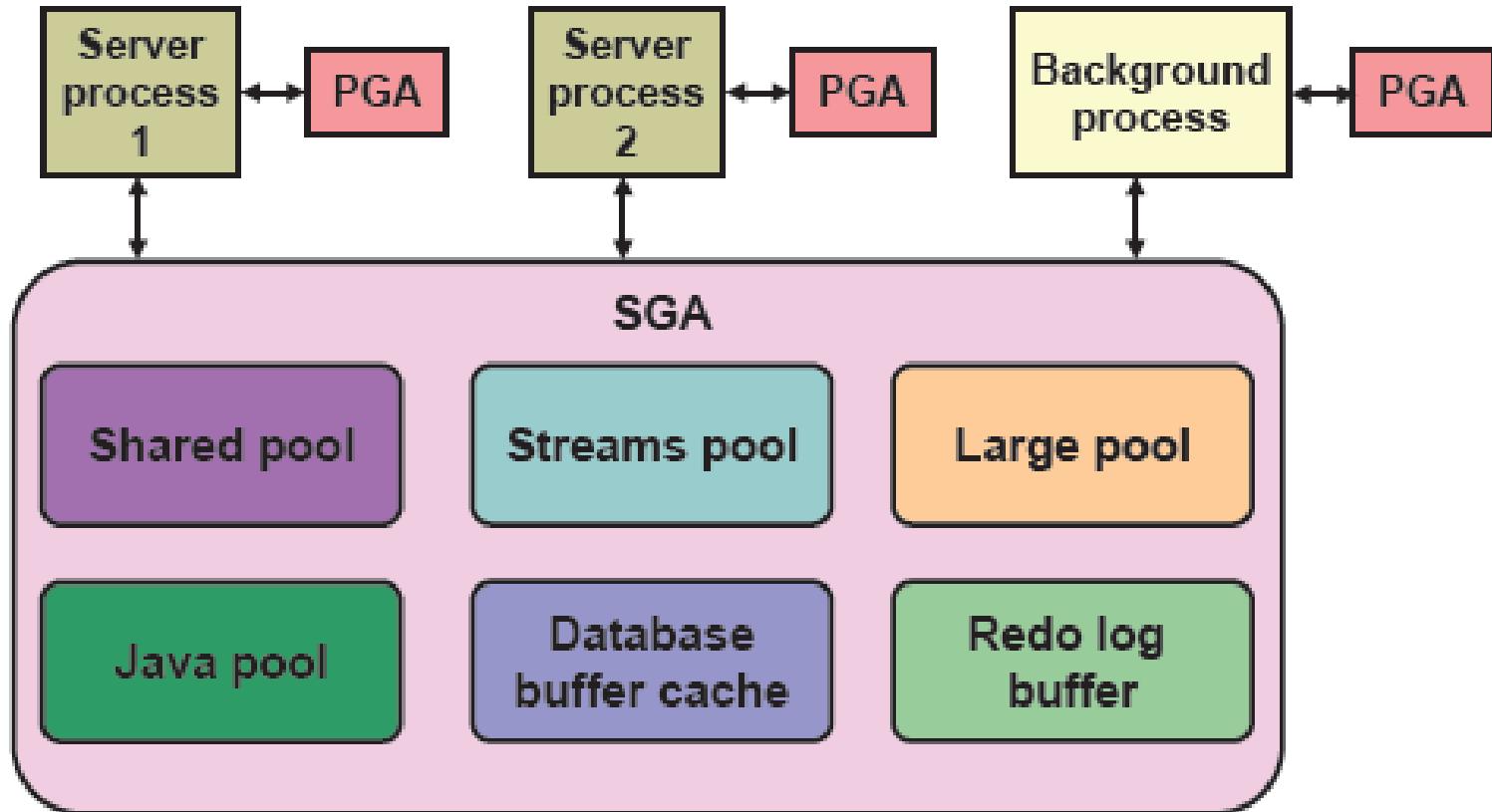
Estructuras de la Base de Datos



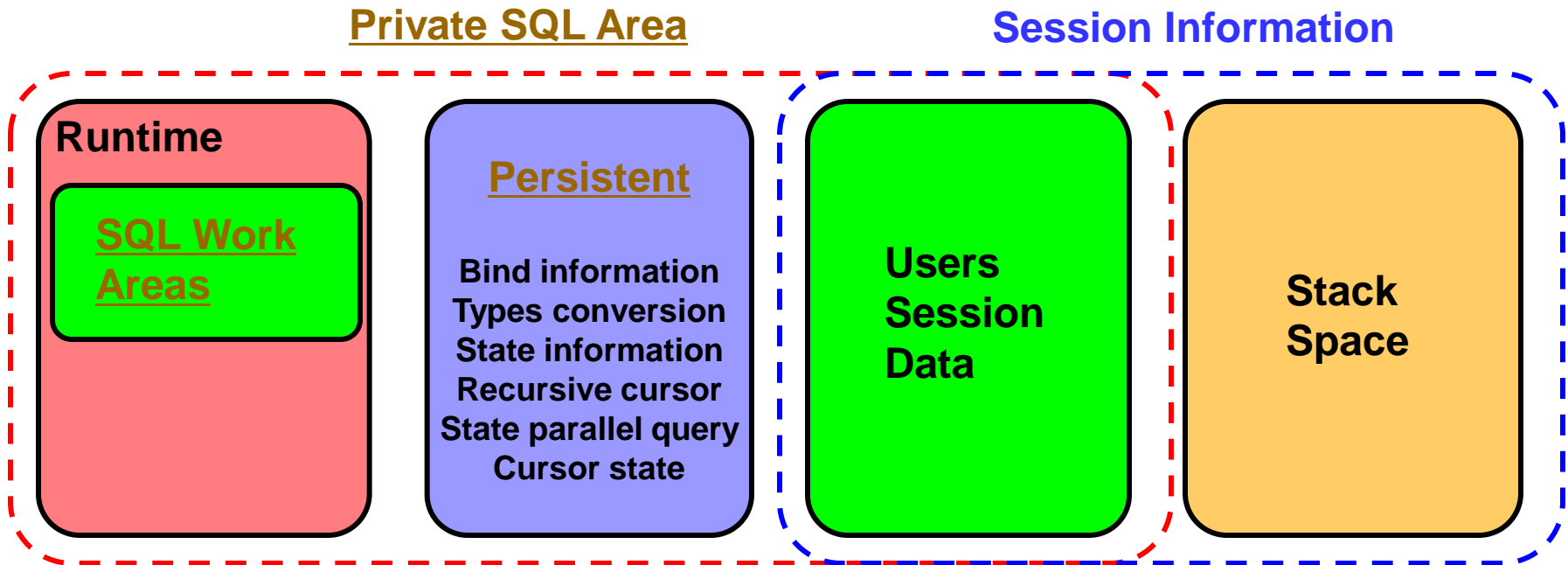
Estructura de la Memoria

- **SGA (Área Global del Sistema):** asignada al iniciar la instancia y componente fundamental de la misma. Es dinámico, se pueden cambiar los tamaños de los distintos componentes sin bajar la instancia
 - **PGA (Área Global de Programas):** asignada al iniciar el proceso servidor. Es privado para cada proceso servidor y proceso background; tiene *stack areas* y *data areas*
-

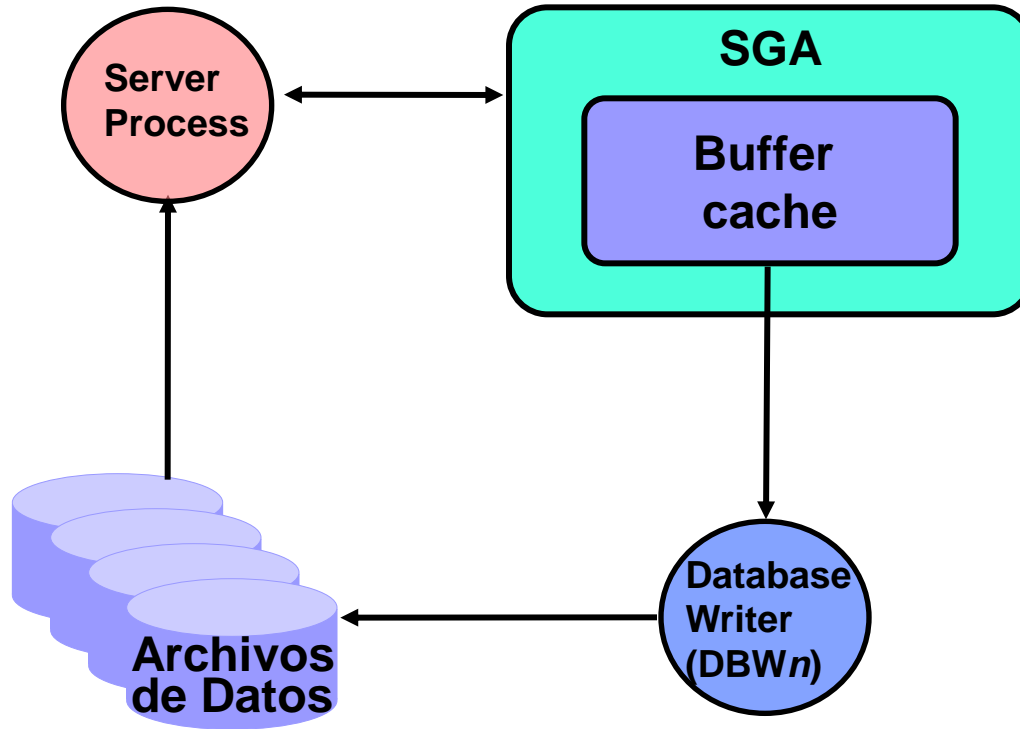
Estructura de la Memoria



PGA Servidores Dedicados



Buffer Cache



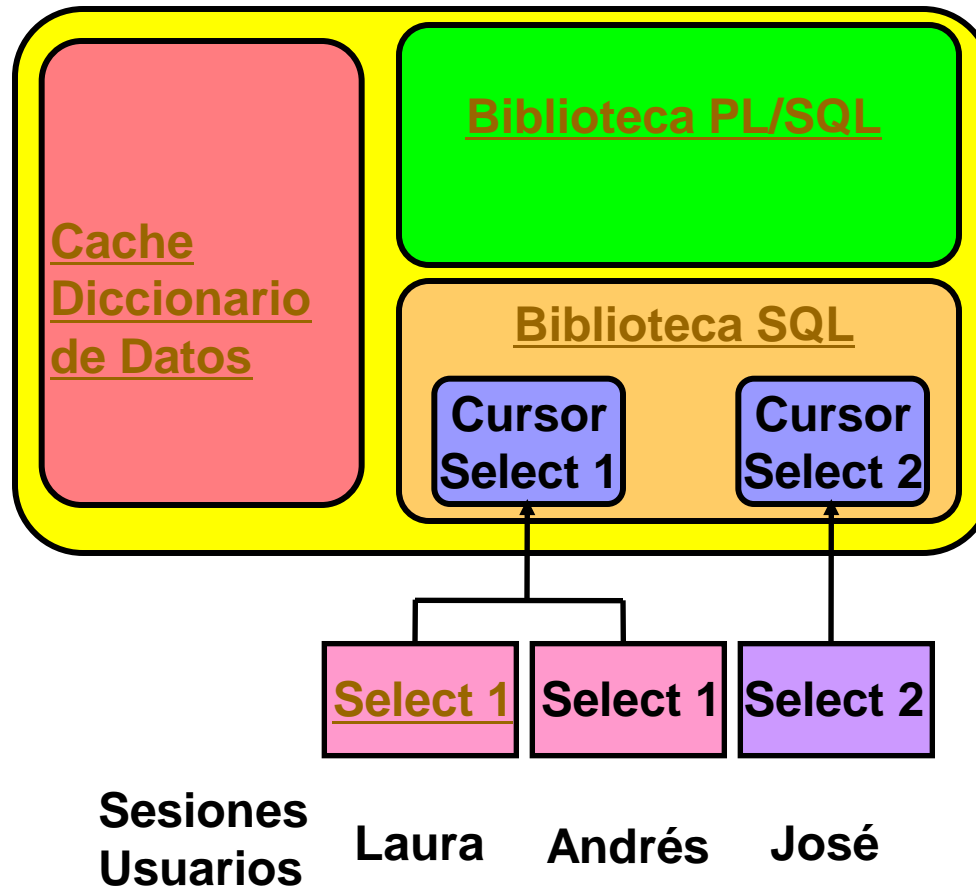
■ Buffers:

- ❑ Sucio
- ❑ Libre o no usado
- ❑ Frío

Se organizan en dos listas:

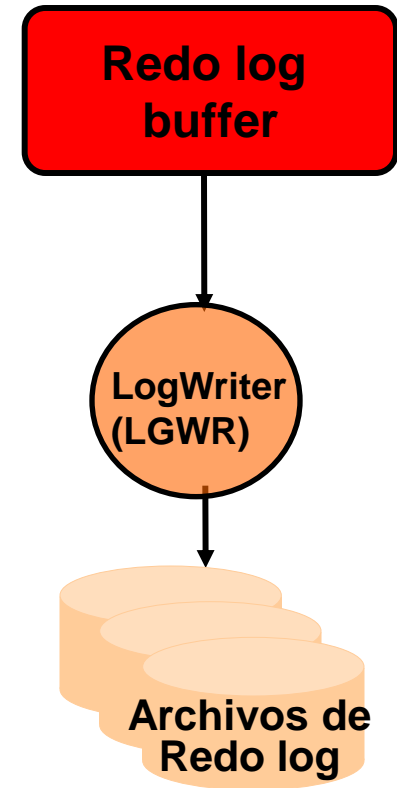
1. Write list: buffers sucios, datos modificados pero aun no escritos a disco
2. LRU list: buffers limpios, buffers “pinned” (accedidos en ese momento) y buffers sucios que aun no están en el Write list

Shared Pool



Redo Log Buffer

- Registra todos los cambios realizados en los bloques de datos
- Estos cambios se denominan registro de redo
- Su objetivo es garantizar la recuperación de la base de datos

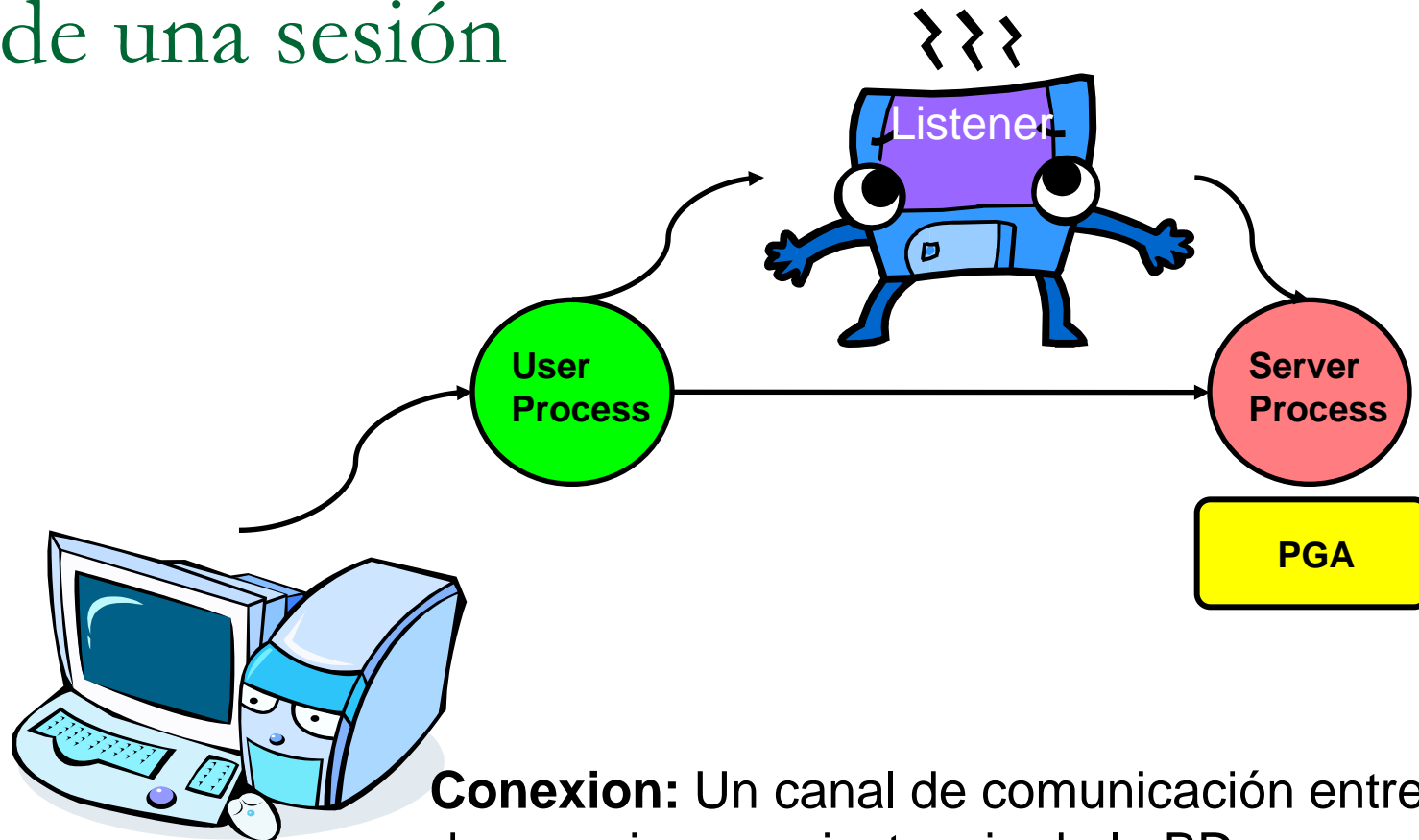


Estructura de los Procesos

Un proceso es un hilo de control o un mecanismo en un sistema operativo que puede ejecutar una serie de pasos

- **Procesos de usuario:** ejecutan la aplicación o código de alguna herramienta. Ej: SQL Plus
 - **Procesos del SMBD:** que ejecutan el código del servidor del SMBD. Esto incluye procesos servidores y de background.
-

Establecimiento de Conexión y creación de una sesión



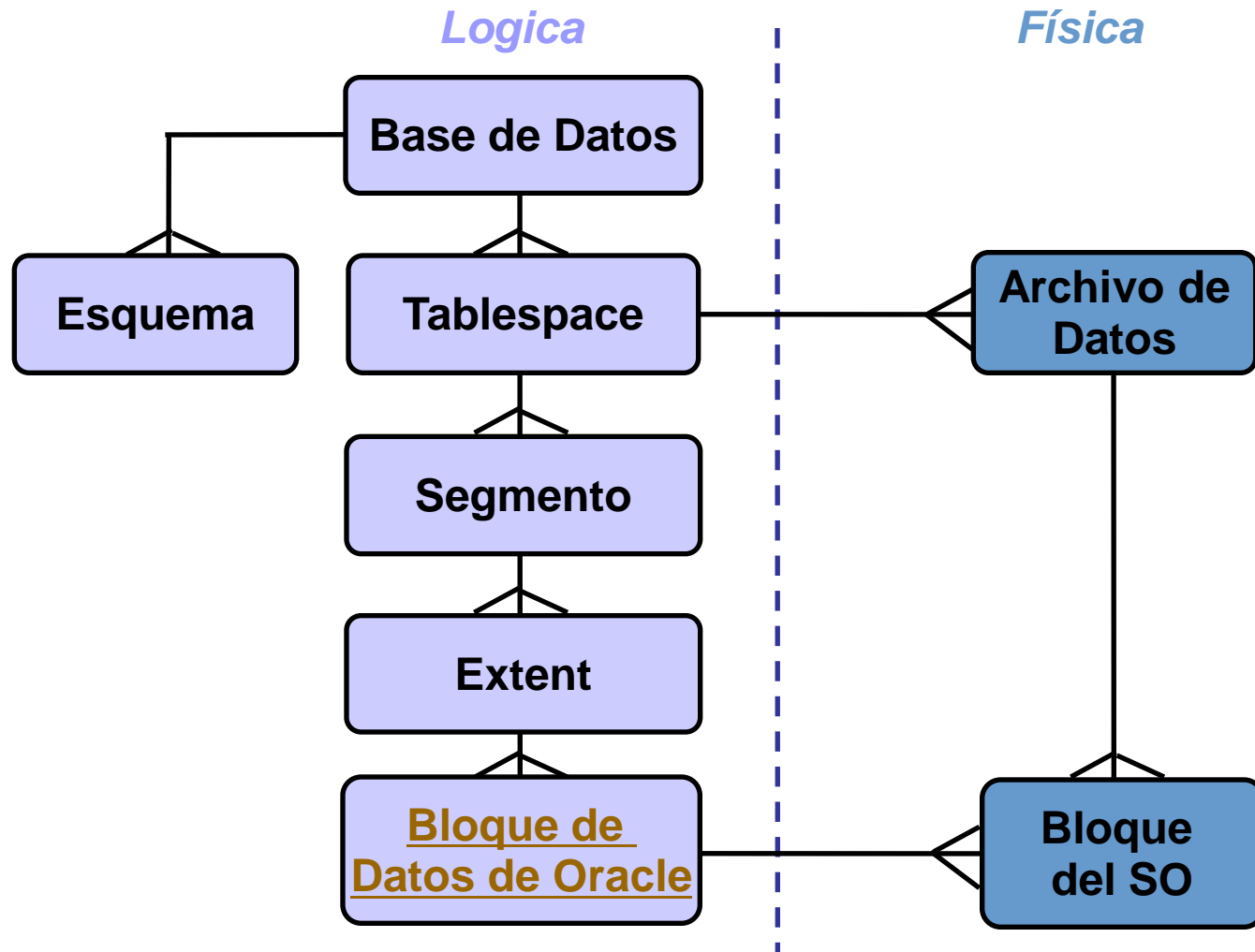
Conexion: Un canal de comunicación entre un proceso de usuario y una instancia de la BD

Sesion: Una conexión específica del usuario a una instancia de la BD a través de un proceso de usuario

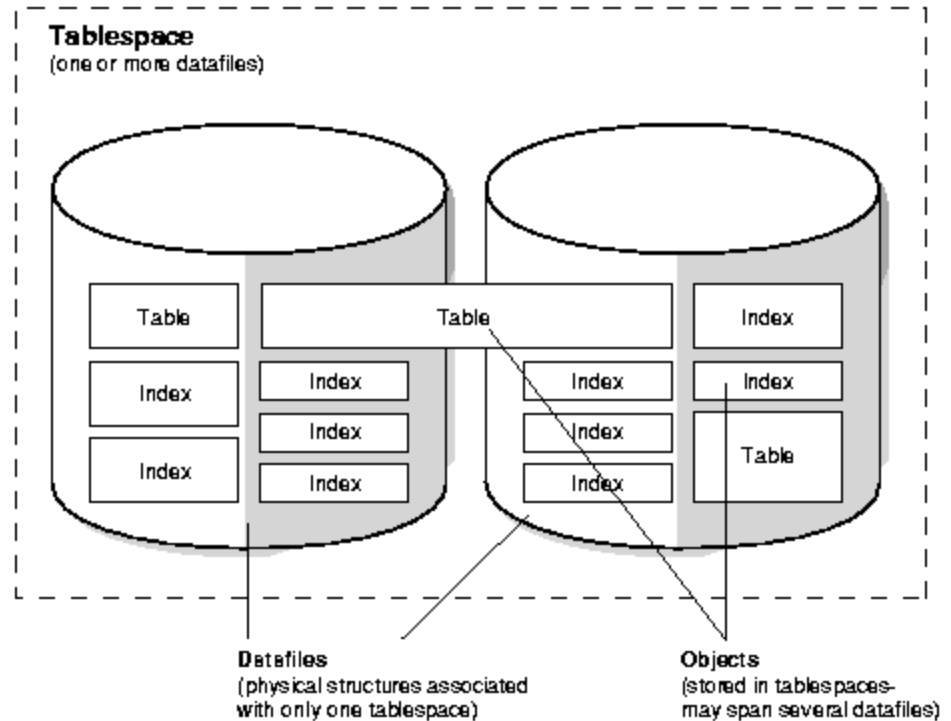
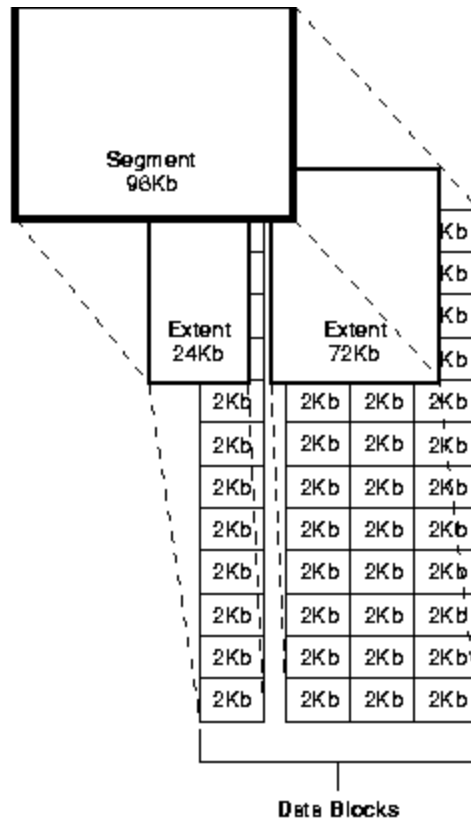
Estructura Física de la Base de Datos

- Archivo de Datos
 - Archivo de Control: datos acerca de la BD
 - Archivos de Redo Log
 - Archivo de Parámetros: configuración instancia
 - Archivos de Respaldos
 - Archivos de Archive Log
 - Archivo de Password
 - Alert Log y
 - Archivos de trazas
-

Relación entre las Estructuras Lógicas y Físicas



Estructuras de Almacenamiento



Diccionario de Datos

- Definiciones de objetos en el esquema
 - Espacio utilizado por los objetos
 - Valores por defecto columnas
 - Restricciones de integridad
 - Usuarios, privilegios y roles
 - Información de auditoría
 - Vistas Estáticas: USER, ALL, DBA
 - Vistas Dinámicas: Ejemplo
-

SGA

- Estructuras de memoria compartida que contienen información de datos y control para una instancia de una base de datos Oracle
- Para un desempeño óptimo el SGA completo debe estar en memoria. Si no lo está y se utiliza la memoria virtual el desempeño puede decrecer dramáticamente a causa de la paginación por parte del sistema operativo del SGA.



Large Pool

- Área opcional que proporciona asignaciones grandes de memoria para ciertos procesos complejos, tales como:
 - Operaciones de respaldo y recuperación
 - Procesos I/O del servidor



Shared Pool

- Contiene las tres áreas mayores: *library cache*, *dictionary cache* y estructuras de control, compartidas por diversos usuarios



Java Pool

- Utilizada por las sesiones de código y datos de Java dentro de la máquina virtual de Java (JVM)



Streams Pool

- Utilizado por Oracle Streams



Database buffer cache

- Mantiene una copia de los bloques leídos de los archivos de datos. Todos los procesos de los usuarios conectados concurrentemente a la instancia compartida lo acceden
- La primera vez que un proceso usuario requiere una pieza de datos, éste busca los datos en el buffer cache. Si el proceso los encuentra se tiene un **cache hit**. En caso contrario debe buscarlos en el archivo de datos y se tiene un **cache miss**



Redo log buffer

- Buffer circular que mantiene información acerca de los cambios hechos en la BD
- Contiene la información necesaria para reconstruir o rehacer, los cambios hechos en la base de datos por las operaciones: INSERT, UPDATE, DELETE, CREATE, ALTER o DROP.
- Hace cache de información de redo (utilizada para recuperar la instancia) hasta que pueda ser escrita en archivos de redo log almacenados en el disco



PMon

- Efectúa proceso de limpieza cuando un proceso de usuario falla
- Limpia el buffer cache de la BD y libera recursos que estaba usando el proceso de usuario



SMon

- Efectúa recuperación, si es necesario, al iniciar una instancia.
- Limpia segmentos temporales que no estén utilizándose.



Procesos servidor

- Manejan los requerimientos de procesos de usuarios conectados a la instancia
 - Análisis y ejecución de instrucciones SQL
 - Leer bloques de datos de datafiles en el disco
 - Devolver resultados para que la aplicación pueda procesar la información



DBW_n

- Escribe el contenido de los buffers (sucios y fríos) a los datafiles en disco
- Administra el buffer cache para que los procesos de usuario siempre encuentren buffers libres



LGWR

- Administra el redo log buffer (buffer circular), escribiendolo a los redo log file en disco.
- Escribe todas las entradas redo que han sido copiadas al buffer desde la última vez
- Antes que DBWn escriba un buffer modificado, todos los registros redo asociados a los cambios del buffer deben ser escritos al disco (the **write-ahead protocol**).



CKPT

- Actualiza los headers de todos los datafiles para registrar los detalles de un checkpoint que ocurra.
- No escribe bloques al disco (esto es siempre efectuado por DBW*n*).



ARCn

- Copia los archivos redo log a un dispositivo de almacenamiento después de ocurrir un log switch.
- Procesos *ARCn* están presentes sólo cuando la BD está en modo ARCHIVELOG, y la archivación automática está habilitada



Buffer cache

- Cuando un buffer en el DB buffer cache es modificado, se marca como sucio.
- Un buffer frío es un buffer que no ha sido recientemente usado de acuerdo al algoritmo menos recientemente usado (LRU)
- Se organizan en dos listas:
 1. Write list: buffers sucios, datos modificados pero aun no escritos a disco
 2. LRU list: buffers limpios, buffers “pinned” (accedidos en ese momento) y buffers sucios que aun no estan en el write list



Online redo log files

- Permiten la recuperación de la instancia de la BD
- Si la BD se cae (*crash*) y no se pierden data files, entonces la instancia puede recuperar la BD con la información de estos archivos



Archive log files

- Contienen la historia continua de los cambios de los datos (redo) generados por la instancia
- Utilizando un backup de la BD y estos archivos se puede recuperar un data file perdido.



Bloques de datos Oracle

- Los datos en la BD son almacenados físicamente en bloques Oracle: la mínima unidad de espacio físico, la cual es un múltiplo del bloque del SO (2 Kb a 52kb).
- El tamaño del bloque Oracle se fija por el parámetro DB_BLOCK_SIZE del fichero init.ora. Un tamaño grande de bloque mejora la eficiencia del cache de E/S, pero el tamaño de la SGA aumentará para contener los mismos DB_BLOCK_BUFFERS, lo que significa un problema de memoria.



Ejecución Select SQL

El proceso cliente pasa la sentencia SELECT SQL al proceso servidor por medio de la SGA.

Los procesos del servidor buscan en la zona de SQL compartido una versión ejecutable de la sentencia. Si la encuentran no tienen que procesarla.

Se procesa la sentencia SQL y su versión ejecutable se coloca en la zona de SQL compartido.



SQL Work Areas

- Para consultas complejas una gran porción del área de ejecución es dedicada a áreas de trabajo asignadas por operadores de memoria intensiva tales como: operadores basados en Sort, Hash-join, Bitmap merge o bitmap create



Área persistente

- Contiene información que debe ser mantenida durante muchas ejecuciones de la misma instrucción por ejemplo, la información de contexto o conversiones de tipos de datos.
- Es liberado sólo cuando el cursor es cerrado.
- Debe tenerse cuidado cuando los cursores se manejan por la aplicación porque se consumen recursos.



Private SQL Areas

- Contiene datos como información vinculada y estructuras de memoria de ejecución
- Cada sesión que efectúa una instrucción SQL tiene un área privada SQL
- Cada usuario que envía la misma instrucción SQL tiene su propia área privada que utiliza un área simple compartida, por lo que diversas áreas privadas pueden ser asociadas con una misma área SQL compartida



Library Caché

- Incluye las áreas de SQL compartidas y privadas en caso de que sea un servidor de múltiples transacciones.
- Los procedimientos PL/SQL y paquetes, estructuras de control como los locks y manejadores del caché de librerías



Dictionary Caché

- Es una colección de tablas y vistas de la base de datos que contienen información de referencia a cerca de la base de datos, sus estructuras y usuarios. Oracle accede los datos del diccionario frecuentemente durante el análisis de una instrucción de SQL. El diccionario de datos es accedido en dos ubicaciones especiales en la memoria designada para éste. Un área es llamada data dictionary cache, también conocida como row caché porque mantiene los datos como tuplas en lugar de buffers (que son bloques de datos enteros). El otra área en memoria para mantener el diccionario es library cache



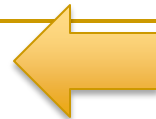
Shared SQL Areas y Private SQL Areas

- Oracle representa la ejecución de cada instrucción SQL con un área SQL privada y otra compartida. Reconoce cuando dos usuarios están ejecutando la misma instrucción SQL y reutiliza el área compartida para dichos usuarios.
- Shared SQL Areas: un área compartida SQL contiene un árbol proveniente del análisis (léxico, sintáctico y semántico) de la instrucción y su plan de ejecución



Diccionario de Datos

- Las tablas del diccionario de datos no se pueden acceder directamente, pero se puede acceder a la información contenida en ellas a través de las vistas del diccionario de datos. Para listar las vistas del diccionario de datos se puede consultar la vista `DICTIONARY`
- Las tablas y vistas son llamadas estáticas, porque ellas cambian sólo cuando un cambio es hecho al diccionario de datos, por ejemplo al crear una nueva tabla o autorizar algún usuario con privilegios.
 - La vista con prefijo `ALL_`, muestra toda la información accesible al usuario actual
 - La vista `DBA_` muestra toda la información relevante de la base de datos entera
 - La vista `USER_` muestra toda la información del esquema del usuario actual



Vistas dinámicas

- Estas vistas monitorean la actividad de la BD son mantenidas por el servidor y accesibles al usuario administrador de la base de datos SYS.
- Ellas son actualizadas continuamente mientras que la base de datos este abierta y en uso y su contenido esta relacionado principalmente al desempeño.
- Aunque estas vistas parezcan ser tablas regulares, no lo son. Las vistas proveen datos en estructuras de disco internas y estructuras de memoria. Un usuario las puede consultar, pero nunca actualizar o alterarlas



Vistas dinámicas: Ejemplo

- Instrucciones SQL y su numero de ejecuciones asociadas donde el tiempo de CPU consumido es mayor a 200.000 microsegundos
 - ❑

```
SELECT sql_text, executions  
FROM v$sql  
WHERE cpu_time > 200000
```

