



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

Asignatura: Bases de Datos Avanzadas

Grupo: 1

Semestre: 2023-1

**Ejercicio Práctico 1: Juego de caracteres y
componentes de la base de datos**

Fecha Límite de Entrega: 12/09/2022

Profesor: Jorge Alberto Rodríguez Campos

Alumna:

❖ Najera Noyola Karla Andrea

Ejercicio Práctico 1: Juego de caracteres y componentes de la base de datos

Objetivo

Comprender la importancia de los juegos de caracteres durante el proceso de instalación de una base de datos, conocer la forma en la que se pueden consultar los diferentes componentes instalados en una base de datos y calcular el espacio en disco que ocupan.

Trabajo previo

C1 - Proporcionar una breve respuesta para los siguientes puntos:

★ ¿Qué significa AL32UTF8?

Es un conjunto de caracteres propio de las bases de datos Oracle que se distingue de UTF-8 debido a que, además de tener los mismos caracteres, tiene soporte para datos de tipo XML.

El "AL" hace referencia a todos los lenguajes (all languages), 32 al tamaño máximo de bits (32 bits=4bytes) y UTF-8 debido a que incluye a todos los caracteres aceptados por este conjunto.

★ ¿Cuál es la longitud máxima que puede tener un carácter con esta configuración?

Al usar AL32UTF8, el tamaño de un carácter puede variar entre 1 y 4 bytes.

★ ¿En qué casos un carácter requeriría la longitud máxima para poder almacenarse?

Cuando se necesiten caracteres adicionales de lenguas asiáticas como Chino, Japonés, Coreano, así como de Lenguas Históricas.

★ ¿Por qué Oracle recomienda este juego de caracteres?

Oracle indica las siguientes ventajas:

- Es un superconjunto de UTF-8
- Es un juego de caracteres universal ya que representa todas las codificaciones de idiomas conocidos.
- Es eficiente en el espacio debido a que utiliza un byte para la mayoría de caracteres de uso común (como ASCII).
- A diferencia de UTF-8, admite datos XML
- Admite valores de 4 bytes, algo que no es posible en UTF-8

- Alta compatibilidad dado las ventajas de Unicode

★ ¿Qué desventajas y en qué situaciones no se recomendaría este valor?

Dado que AL32UTF8 es un conjunto de caracteres de tamaño variable, las operaciones en la base de datos en aquellos datos que contemplen caracteres pueden llegar a ser un poco más lentas en comparación con los conjuntos de caracteres de la base de datos que utilicen un solo byte, como lo son WE8ISO8859P1 o WE8MSWIN1252. Asimismo, porque los espacios requeridos para su almacenamiento son más altos en comparación de aquellos que solo usan ASCII y, por lo mismo, datos que se encuentren en inglés pueden requerir más espacio si se almacenan en columnas CLOB.

Por ello, podría optarse por modificar este valor por defecto si tenemos algún entorno en donde solo usemos datos nacionales que no requieren del almacenamiento de caracteres especiales, así como de aquellos sistemas que no tengan altos requerimientos de compatibilidad con otros conjuntos de caracteres.

Desarrollo

C2 - Código del script s-01-database-info.sql debidamente formateado

```
--@Autor:      Najera Noyola Karla Andrea
--@Fecha creación: 07/09/2022
--@Descripción:  Creación de tabla con información de la base de datos

--esta instrucción permite detener la ejecución del script al primer error
--útil para detectar errores de forma rápida.
--al final del script se debe invocar a whenever sqlerror continue none
--para regresar a la configuración original.
whenever sqlerror exit rollback;

-- para propósitos de pruebas y propósitos académicos se incluye password
-- no hacer esto en sistemas reales.
Prompt conectando como usuario sys
connect sys/system1 as sysdba
declare
    v_count number;
    v_username varchar2(20) := 'KARLA0201';
begin
    select count(*) into v_count from all_users where username=v_username;
    if v_count >0 then
        execute immediate 'drop user '||v_username|| 'cascade';
    end if;
end;
/
```

Prompt creando al usuario karla0201

```
create user karla0201 identified by karla quota unlimited on users;  
grant create session, create table to karla0201;
```

Prompt creando tabla informativa de la base de datos

```
create table karla0201.database_info(  
    instance_name varchar2(16),  
    db_domain varchar2(20),  
    db_charset varchar2(15),  
    sys_timestamp varchar2(40),  
    timezone_offset varchar2(10),  
    db_block_size_bytes number(5,0),  
    os_block_size_bytes number(5,0),  
    redo_block_size_bytes number(5,0),  
    total_components number(5,0),  
    total_components_mb number(10,2),  
    max_component_name varchar2(30),  
    max_component_desc varchar2(64),  
    max_component_mb number(10,0)  
);
```

Prompt insercion de valores

```
insert into karla0201.database_info(instance_name, db_domain, db_charset,  
sys_timestamp,  
    timezone_offset, db_block_size_bytes, os_block_size_bytes,  
redo_block_size_bytes,  
    total_components, total_components_mb, max_component_name,  
max_component_desc,  
    max_component_mb) values (  
    --instance_name  
    (select instance_name from v$instance),  
    --db_domain  
    (select value from v$parameter where name='db_domain'),  
    --db_charset  
    (select value from nls_database_parameters where  
parameter='NLS_CHARACTERSET'),  
    --sys_timestamp  
    (select systimestamp from dual),  
    --timezone_offset  
    (select tz_offset((select sessiontimezone from dual)) from dual),  
    --db_block_size_bytes  
    (select value from v$parameter where name='db_block_size'),  
    --os_block_size_bytes  
    (512),  
    --redo_block_size_bytes  
    (select blocksize from v$log where group#=1),  
    --total_components  
    (select count(*) from v$sysaux_occupants),  
    --total components_mb
```

```

        (select round((sum(space_usage_kbytes)/1024),2) from v$sysaux_occupants),
--max_component_name
        (select occupant_name from v$sysaux_occupants where space_usage_kbytes=
        (select max(space_usage_kbytes) from v$sysaux_occupants)),
--max_component_desc
        (select occupant_desc from v$sysaux_occupants where space_usage_kbytes=
        (select max(space_usage_kbytes) from v$sysaux_occupants)),
--max_component_mb
        (select round((space_usage_kbytes/1024), 2)
        from v$sysaux_occupants where space_usage_kbytes=
        (select max(space_usage_kbytes) from v$sysaux_occupants))
);

```

Prompt mostrando datos parte 1

```
set linesize window
```

```
select instance_name,db_domain,db_charset,sys_timestamp,timezone_offset
from karla0201.database_info;
```

Prompt mostrando datos parte 2

```
select db_block_size_bytes,os_block_size_bytes,redo_block_size_bytes,
total_components,total_components_mb
from karla0201.database_info;
```

Prompt mostrando datos parte 3;

```
select max_component_name,max_component_desc,max_component_mb
from karla0201.database_info;
```

```
whenever sqlerror continue none;
```

C3 - Salida de ejecución del script anterior.

```

Terminal - karla@pc-knn:/unam-bda/Ejercicios/Tema2/Ej1
File Edit View Terminal Tabs Help
===> Eliminando al usuario karla0201
Invocando scripts del ejercicio
conectando como usuario sys
Connected.
creando al usuario karla0201
creando tabla informativa de la base de datos
inserción de valores
mostrando datos parte 1

```

INSTANCE_NAME	DB_DOMAIN	DB_CHARSET	SYS_TIMESTAMP	TIMEZONE_0
knnbda1	fi.unam	AL32UTF8	07-SEP-22 05.30.47.973331 PM -05:00	-05:00

```

mostrando datos parte 2

```

DB_BLOCK_SIZE_BYTES	OS_BLOCK_SIZE_BYTES	REDO_BLOCK_SIZE_BYTES	TOTAL_COMPONENTS	TOTAL_COMPONENTS_MB
8192	512	512	32	475.88

```

mostrando datos parte 3

```

MAX_COMPONENT_NAME	MAX_COMPONENT_DESC	MAX_COMPONENT_MB
SDO	Oracle Spatial	196

```

Script invocado. Presionar ENTER para comenzar la validación o Ctrl +C para cancelar.

```

C4- Salida de ejecución del validador.

```
Terminal - karla@pc-knn:/unam-bda/Ejercicios/Tema2/Ej1
File Edit View Terminal Tabs Help
=====
Iniciando proceso de validación de respuestas
Incluir en el reporte a partir de este punto
=====
SESSION_TIME: 07/09/2022 17:30:55
USR_COD_HOST: 715009
OS_USER: karla
BD_USER: SYS
CON_NAME: knnbda1
8db764d9cc93705ca30ee920c575812c81abc63171219ab5d3fe0d766fb74a35 s-02p-validador-ejercicios.plb

2022-09-07 17:30:56.175-715009-PY.K001K.Y001S.Y001 ==> OK Tabla database_info encontrada
2022-09-07 17:30:56.177-715009-CV.A002N.V002Y.V002 ==> OK 1 registro encontrado en karla0201.database_info
2022-09-07 17:30:56.179-715009-OR.R003N.R003S.R003 ==> OK Nombre de instancia valido: knnbda1
2022-09-07 17:30:56.180-715009-KO.L004B.O004S.O004 ==> OK Dominio correcto: fi.unam
2022-09-07 17:30:56.181-715009-NG.A005D.G005Y.G005 ==> OK Juego de caracteres correcto: AL32UTF8
2022-09-07 17:30:56.181-715009-NQ.K006A.Q006S.Q006 ==> OK sys timestamp: 07-SEP-22 05:30:47.973331 PM -05:00
2022-09-07 17:30:56.181-715009-OC.A007I.C007S.C007 ==> OK Offset de la zona horaria correcto: -05:00
2022-09-07 17:30:56.183-715009-FP.R008K.P008Y.P008 ==> OK Tamaño de bloque para BD correcto: 8192
2022-09-07 17:30:56.184-715009-IQ.L009N.Q009S.Q009 ==> OK Tamaño de bloque para s.o. correcto: 512
2022-09-07 17:30:56.185-715009-OW.A010N.W010S.W010 ==> OK Tamaño de bloque para Redo Logs correcto: 512
2022-09-07 17:30:56.185-715009-UT.K011B.T011Y.T011 ==> OK Número de componentes instalados: 32
2022-09-07 17:30:56.186-715009-NS.A012D.S012S.S012 ==> OK Espacio en MB que ocupan los componentes: 475.88
2022-09-07 17:30:56.186-715009-AB.R013A.B013S.B013 ==> OK Componente con el mayor espacio: SDO
2022-09-07 17:30:56.187-715009-MQ.L014I.Q014Y.Q014 ==> OK Descripción del componente: Oracle Spatial
2022-09-07 17:30:56.187-715009-PX.A015K.X015S.X015 ==> OK Espacio en MB del componente mayor: 196
2022-09-07 17:30:56.188-715009-CX.K016N.X016S.X016 ==> OK Validación concluida.
Removiendo SPV_ASSERT
Removiendo SPV_CREA_SECUENCIA
Removiendo SPV_ELIMINA_USUARIO
Removiendo SPV_PRINT_HEADER
Removiendo SPV_PRINT_OK
Removiendo SPV_THROW_ERROR
Removiendo SPV_VALIDA_DB_INFO
Removiendo FV_HASH
Disconnected from Oracle Database 19c Enterprise Edition Release 19.0.0.0.0 - Production
Version 19.3.0.0.0
[karla@pc-knn Ej1]$
```

Conclusiones y comentarios

En este ejercicio me fue posible entender un poco más de las configuraciones que realizamos al momento de instalar de una base de datos ya que, por quedarnos con los valores por default que proporcionan los instaladores gráficos como dbca, no conocemos realmente las implicaciones que tienen al momento de realizar inserciones en la base de datos y, en el caso del juego de caracteres cobra mayor importancia debido a que colocar algún conjunto que no sea adecuado a los requerimientos funcionales de un determinados sistema puede generar la aparición de inconsistencias e incompatibilidades.

Por lo anterior, considero que fue importante conocer las características que tenemos por default en la primera base que creamos con el fin de ponerle mayor atención en la generación de futuras bases de datos.

Referencias

Campana, D. (2022, 22 marzo). *Base de datos en memoria TimesTen*. Oracle Blogs.
<https://blogs.oracle.com/timesten/post/why-databasecharacter-set-matters>

Character set when upgrade from Oracle 11.2.0.4 to 12.2. (2019). Oracle Communities.
<https://community.oracle.com/tech/developers/discussion/4281084/character-set-when-upgrade-from-oracle-11-2-0-4-to-12-2>

Oracle. (2005, mayo). *Globalization Support- Oracle Unicode Database Support*.
<https://www.oracle.com/technetwork/products/globalization/twp-appdev-unicode-10gr2-129234.pdf>