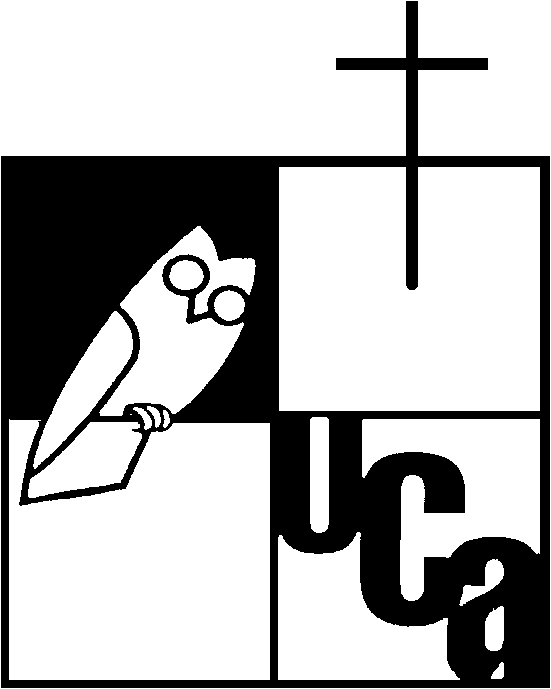
Universidad Centroamericana “José Simeón Cañas”.

Ingeniería Informática.

Administración de bases de datos.

Exposición de bases de datos.

Integrantes:

* Fredy Alexander Sánchez Pérez Carnet: 00082817
* Carlos Mauricio Paredes Pineda Carnet: 0028516
* Mario Andrés Muñoz Kessler Carnet: 00405016
* Fernando Andrés Quezada Menjívar Carnet:00195316
* Reynaldo Alexander Cordova Quintana Carnet: 00404416

Catedrático: Lic. Erick Giovanni Varela.

Fecha de entrega: jueves, 08 de noviembre de 2018

Índice

**Introducción. 2**

[**Marco teórico:**](#_3hjqqoih3prd) **3**

[CLUSTERS](#_ep3mqf888zgh) 3

[Prerrequisito para crear un clusters](#_hvo0emphusto)

[Sintaxis para crear un clusters:](#_os9oehnl7kx) 4

[Restricciones de datos](#_k02fwp15dvgl) 4

[Parámetros de clusters:](#_68f4gi9nu372)

Cláusula de atributos físicos 5

Cláusula de índice 5

Cláusula HASHKEYS 5

Norowdependencies | Rowdependencies 5

CACHE | NOCHE 6

TAMAÑO 6

**Conclusiones 7**

**Bibliografía 8**

**Anexos**

Cuestionario 9

# Introducción.

Mediante este trabajo estamos dando a conocer que es el de tema de la administración de clusters. Las empresas utilizan la tecnología para obtener una ventaja competitiva, reducir los costos operativos, mejorar la comunicación con los clientes y aumentar la visibilidad de administración de los procesos.

Nuestro grupo de trabajo investigó sobre el tema de clusters para ver la funcionalidad que tiene este sobre el sistema y así conocer la herramienta, para después explicarle a nuestros compañeros.

Marco teórico:

Clusters

Un clúster proporciona un método opcional para almacenar datos de tablas. Un clúster se construye a partir de un grupo de tablas que comparten los mismos bloques de datos. Las tablas se agrupan porque comparten columnas comunes [1].

Debido a que los clústeres almacenan filas relacionadas de diferentes tablas juntas en los mismos bloques de datos, los clústeres utilizados correctamente ofrecen dos ventajas principales:

- La Entrada / Salida de disco se reduce y el tiempo de acceso mejora para las combinaciones de tablas agrupadas [1].

- La **clave de clúster** es la columna, o grupo de columnas, que las tablas agrupadas tienen en común.Se debe especificar las columnas de la clave de clúster al crear el clúster. Cada valor de clave de clúster se almacena sólo una vez en el clúster y el índice del clúster, sin importar cuántas filas de diferentes tablas contengan el valor [1]

Después de crear un clúster, puede crear tablas en el clúster. Sin embargo, antes de que se puedan insertar filas en las tablas agrupadas, se debe crear un índice de agrupación. El uso de clústeres no afecta la creación de índices adicionales en las tablas agrupadas. Se pueden crear y soltar como de costumbre [1].

Prerrequisito para crear un clusters

Para crear un clúster en su propio esquema, debe tener el privilegio de sistema CREATE CLUSTER. Para crear un clúster en el esquema de otro usuario, debe tener CREATE ANY CLUSTER privilegios de sistema. Además, el propietario del esquema para contener el clúster debe tener una cuota de espacio en el espacio de tablas que contiene el clúster o el UNLIMITED TABLESPACE[2].

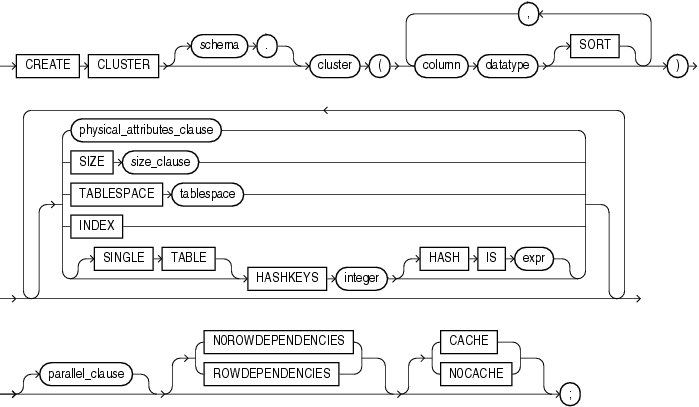
Oracle Database no crea automáticamente un índice para un clúster cuando el clúster se crea inicialmente. Las declaraciones en lenguaje de manipulación de datos (DML) no se pueden emitir contra tablas de clúster en un clúster indexado hasta que cree un índice de clúster con una CREATE INDEX declaración [2].

### 

### 

### 

### Sintaxis para crear un clusters:



Semántica.

**Schema**: Especifique el esquema para contener el clúster. Si omite schema, la base de datos Oracle crea el clúster en su esquema actual.

**Cluster**: Especifica el nombre del clúster que se creará.

**Column**: Especifica uno o más nombres de columnas en la clave de clúster.

**Datatype**: Especifica el tipo de datos de cada columna de clave de clúster.

Restricciones de datos

Los clusters tienen restricciones en los tipos de datos que son:

No se puede especificar un clúster columna de clave de tipo de datos LONG, LONG RAW,REF , tabla anidada, ARRAY,BLOB ,CLOB,BFILE, o tipo de objeto definido por el usuario [2].

No puede utilizar la HASH IS cláusula si algún tipo de datos de columna no lo es INTEGER o NUMBER con escala 0 [2].

Puede especificar una columna de tipo ROWID, pero Oracle Database no garantiza que los valores en dichas columnas sean filas válidas [2].

SORT

La palabra clave SORT es válida solo si está creando un clúster hash. Esta cláusula indica a la base de datos Oracle que ordene las filas del clúster en esta columna antes de aplicar la función hash. Si lo hace, puede mejorar el tiempo de respuesta durante las operaciones posteriores en los datos agrupados [2].

Cláusula de atributos físicos

Él PHYSICAL\_ATTRIBUTES\_CLAUSE permite especificar las características de almacenamiento del clúster. Cada tabla del clúster también utiliza estas características de almacenamiento. Si no especifica valores para estos parámetros, Oracle Database utiliza los siguientes valores predeterminados:

-PCTFREE: 10

-PCTUSED: 4

- INITRANS: 2 o el valor predeterminado del espacio de tablas para contener el clúster, el que sea mayor [2].

SIZE

Especifica la cantidad de espacio en bytes reservados para almacenar todas las filas con el mismo valor de clave de clúster o el mismo valor de hash. Este espacio determina el número máximo de valores de cluster o hash almacenados en un bloque de datos.

Sí SIZE no es un divisor del tamaño del bloque de datos, la base de datos Oracle utiliza el siguiente divisor más grande. Sí SIZE es más grande que el tamaño del bloque de datos, entonces la base de datos usa el tamaño del bloque del sistema operativo, reservando al menos un bloque de datos para cada cluster o valor hash.

Cláusula de índice (Index)

Para crear un cluster indexado. En cluster indexado, la base de datos Oracle almacena juntas filas que tienen el mismo valor de clave de clúster. Cada valor de clave de cluster distinto se almacena sólo una vez en cada bloque de datos, independientemente del número de tablas y filas en que se produzca. Si no especifica ni INDEX tampoco HASHKEYS, a continuación, la base de datos Oracle crea un cluster indexada por defecto.

Después de crear un cluster indexado, debe crear un índice en la clave del cluster antes de poder emitir cualquier declaración en lenguaje de manipulación de datos (DML) en una tabla del cluster. Este índice se llama el índice de cluster

Cláusula HASHKEYS

Cláusula para crear un cluster hash y especifique el número de valores hash para el cluster hash. En un cluster hash, la base de datos Oracle almacena filas que tienen el mismo valor de clave hash. El valor hash para una fila es el valor devuelto por la función hash del cluster.

La base de datos Oracle redondea el valor HASH KEYS al número primos más cercano para obtener el número real de valor hash. El valor mínimo para este parámetro es 2. Si omite la INDEX cláusula y el HASHKEYS parámetro, la base de datos crea un clusters indexado de forma predeterminada.

*parallel\_clause.*

El parallel\_clause permite paralelizar la creación del clúster.

Norowdependencies | Rowdependencies

Esta cláusula le permite especificar si CLUSTER utilizara el seguimiento de dependencias a nivel de fila. Con esta función, cada fila en las tablas que conforman el clusters tiene un número de cambio de cambio de sistema (SCN) que representa un tiempo mayor o igual al tiempo de confirmación de la última transacción que modificó la fila.

- Rowdependencies: Especificar esta opción si se desea habilitar el seguimiento de dependencias a nivel de fila. Esta configuración es útil principalmente para permitir la propagación paralela en entornos de replicación. Y esta opción el tamaño de fila en 6 bytes.

- Norowdependencies: Si esta opción no activa la función de dependencias a nivel de fila. Este es el valor por predeterminado

CACHE | NOCHE

-Cache: si desea que los bloques recuperados para este clúster se coloquen en el extremo más recientemente utilizado de la lista menos recientemente utilizada (LRU) en el caché del búfer cuando se realiza un escaneo completo de la tabla. Esta cláusula es útil para tablas de búsqueda pequeñas.

-Nocache: si desea que los bloques recuperados para este clúster se coloquen en el extremo de la lista LRU menos recientemente utilizado en la memoria caché del búfer cuando se realice un escaneo completo de la tabla. Este es el comportamiento predeterminado.

Clustered Tables

El beneficio de crear clusters de tablas es que agiliza el proceso de lectur de datos, ya que almacena todos los datos de una tabla dentro de un mismo bloque(generalmente cada tabla se almacena dentro de un bloque, y otras tablas en otros bloques). Otro beneficio de crear clusters es que reduce la necesidad de repetir valores idénticos a través de diferentes tablas, por ejemplo: si tenemos la tabla producto y la tabla ventas y ambas comparten product\_id, un cluster almacenaría todos los atributos de la tabla producto y todos los atributos de la tabla ventas unidos en una sola fila y compartiendo un único product\_id a diferencia que cuando estaban las tablas separadas, cada una tenía su propia columna de product\_id.

Viéndolo con una analogía digamos que un departamento de recursos humanos tiene dos cajones: uno de empleados y otro de departamentos. Buscar un empleado solo implica ir a ese cajón y buscar el empleado, de la misma forma buscar un departamento solo se va a ese cajón y se saca el folder perteneciente al departamento. Pero si constantemente me piden buscar el folder departamento y los folders de todos los empleados que pertenecen a ese departamento, esto me obliga a buscar primero en cajón de departamento y luego ir a buscar al otro cajón revisando todos los empleados que coinciden a dicho departamento. Una solución fuera tener un solo cajón en el cual los folderes van separados en subcajas donde cada subcaja(o división) contiene todos los empleados de un departamento y el folder del departamento, así se vuelve mucho más fácil recuperar la información sobre los empleados de un específico departamento ya que no hay que buscarlos en otra caja, si no solo se saca directamente ese sub cajón. De la misma forma trabajan los clusters, unen información relevante entre distintas tablas y las almacenan juntas, con el fin de agilizar la obtención de datos.

Los clusters tienen los siguientes beneficios:

* La lectura de datos es reducida para joins entre clustered tables
* el tiempo para accesar dichas tablas se disminuye
* se requiere de menos espacio para guardar tablas relacionadas ya que el índice que une ambas tablas se es guardado una única vez.

No se recomienda hacer clustered tables si se tiene alguna de las siguientes situaciones:

* Si las tablas son muy constantemente actualizadas
* si las tablas frecuentemente requieren un full table scan (recuperar secuencialmente todas las filas de una tabla).
* Las tablas deben ser truncadas.

Hay dos formas de crear clustered tables, por index y por hash.

Los Indexed Cluster tables funcionan creando un index key basado en la columna que tienen en común en la table. Esto lo que crea es un árbol binario (en este caso un heap) y cada fila de las columnas se accede usando dicho index.

Los Hash cluster tables funcionan utilizando un número para crear una hash function y almacenar los datos. Lo que hace es aparte de también tener todas las tablas dentro de un mismo bloque de memoria, lo que hace es que separa todas las filas en diferentes “baldes” basados en haciendo hashing del valor de la columna en común. Estos son beneficiosos para hacer búsquedas rápidas, como por ejemplo buscar un empleado con el id = 3, en vez de ir buscando por toda la tabla, este hace hash(3) y esta función devuelve la ubicación exacta en memoria donde esta esa fila, o al menos el contenedor (que contiene pocos valores) en donde puede encontrar ese empleado. Es malo para cuando se está buscando rangos de información (<,>, <=, >=).

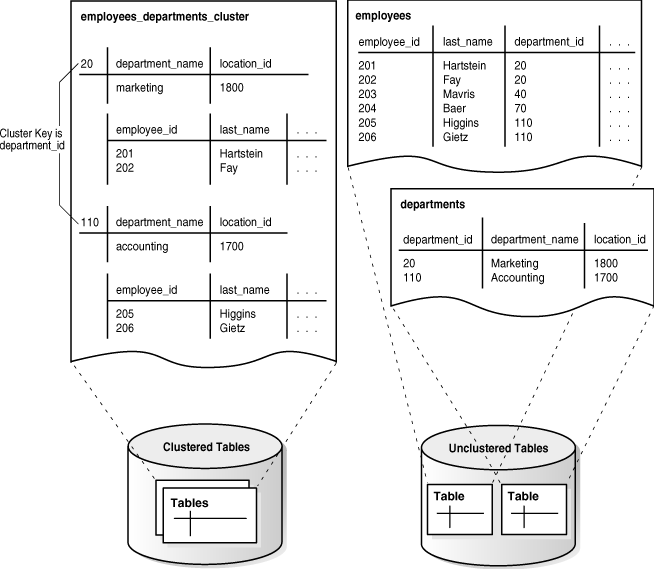
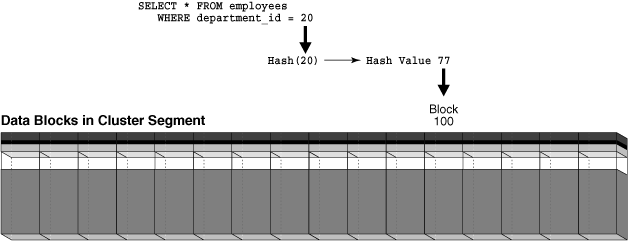


Diagrama de como se ven las Indexed clustered tables vs tablas normales



Acceso de un HASH Cluster

CLUSTER INDEXES

Un índice es un objeto de usuario que contiene una entrada por cada valor que aparece en la columna de índice de la tabla a Cluster, el cual provee un acceso **directo** y **rápido** hacia las columnas  
  
Para crear un Cluster Index, se debe cumplir una de las siguiente condiciones:

-El usuario en cuestión contiene el cluster.

-El usuario en cuestión tiene el privilegio de sistema CREATE ANY INDEX .  
  
Además se debe poseer una quota de tablespace suficiente para contener el Cluster Index, o el privilegio de sistema UNLIMITED TABLESPACE;

Notese que despues de crear un Cluster, se pueden crear tablas en el Cluster pero antes que cualquier fila sea insertada en las Clustered Tables, un Cluster Index se debe crear.  
Los índices en general pueden ser unique, sin embargo los Cluster Indexes no pueden ser unique.  
  
CREATE INDEX emp\_dept\_index \*  
 ON CLUSTER emp\_dept \*  
 TABLESPACE users  
 STORAGE (INITIAL 50K  
 NEXT 50K  
 MINEXTENTS 2  
 MAXEXTENTS 10  
 PCTINCREASE 33);

(\*Campos obligatorios)

Al crear un Cluster Index, los campos obligatorios son el nombre del índice y el cluster al cual pertenece.  
  
Luego indicamos el tablespace que utilizará, seguido de las especificaciones de almacenamiento.  
  
En las especificaciones de almacenamiento asignamos un valor inicial de almacenamiento con INITIAL, luego asignamos el tamaño de la próxima extensión que se guardará en el objeto con NEXT .   
Seguido de esto, asignamos el número mínimo y máximo de extensiones permitidas con MINEXTENTS y MAXEXTENTS.  
  
Por último asignamos el porcentaje de crecimiento que la tercera y demás extensiones tomarán sobre la extensión precedente con PCTINCREASE, su valor por defecto es 50, lo que implica que cada una de las siguiente extensiones es 50% más larga que su precedente. Su valor mínimo es 0, lo cual significa que todas las extensiones serán del mismo tamaño.  
Los índices en general pueden ser unique, sin embargo los Cluster Indexes no pueden ser unique, tampoco puede tener una columna definida como long.

Conclusiones:

1. Respecto a la investigación realizada en este trabajo podemos ver como se puede utilizar una herramienta bastante funcional que son los clusters ayuda a la efectividad del sistema. Bajando los tiempos de respuesta del servidor donde estos están implementados.
2. En este documento podemos ver los atributos que estos contienen la sintaxis de la herramienta de clusters como afecta todo eso en la funcionalidad del clusters y así afectar todo el rendimiento del todo sistema
3. Lo que se gana con este tipo de trabajos que nosotros por nuestra cuenta trabajemos por cuenta propia como estudiante de la carrera de Ing.Informática se nos quiere enseñar a trabajar por nuestra propia cuenta y llegar a tener más información y así jamás de aprender

Bibliografia:

|  |  |
| --- | --- |
| [1] | ORACLE, "Database SQL Reference," ORACLE, [Online]. Available: https://docs.oracle.com/cd/B19306\_01/server.102/b14200/statements\_5001.htm. [Accessed 4 11 2018]. |
| [2] | ORACLE, "ORACLE," ORACLE, [Online]. Available: https://docs.oracle.com/cd/B28359\_01/server.111/b28286/statements\_5001.htm#SQLRF01201. [Accessed 4 11 2018]. |

Anexos:

Cuestionario:

1)¿El parámetro SORT es válido para que tipo de clusters?

a.Índice b.Hash c.CACHE d.CLOB

2) ¿Máximo de tablas que puede contener un cluster?

a.No hay límite b. 6 c.32 c. 24

3)¿Cuál es el máximo de columnas que se pueden especificar como llaves o claves custer?

a.16 columnas b.512 columnas c.32 columnas d.8 columnas

4)¿Cuál o cuáles de las siguientes restricciones tienen los Cluster Iindexes?

a.Incluir una columna definida como LONG

b..Incluir una columna definida como RAW

c.Poseer el atributo NOT NULL

d.Poseer el atributo UNIQUE

5)¿Que no garantiza oracle si se pone el parámetro ROWID?

a.filas válidas b.columnas no válidas c.filas no validas d. Ninguna de la anteriores

6) ¿El parámetro sort en qué momento se activa con respeto a la función hash?

a.Al mismo tiempo b.Después c.Antes d.Ningún momento

7)¿ Para crear un clusters en otro esquema qué privilegio necesita?

a.CREATE ANY CLUSTERS b.CREATE CLUSTERS c.CLUSTERS CREATE d.CREATE

8) ¿Qué sucede si no se especifica la clave INDEX al crear un Cluster?

A: NO SE CREA EL CLUSTER.

B: EL CLUSTER SE CREA Y ADEMÁS SE CREA UN CLUSTER INDEX POR DEFECTO.

C: EL CLUSTER SEA CREA PERO CON ERRORES.

D: EL CLUSTER ESTALLA

9)¿Cuál o cuáles de las siguientes condiciones se deben cumplir para crear un Cluster Index?

A)El usuario en cuestión contiene el cluster.

B) El usuario en cuestión tiene el priviliegio de sistema CREATE ANY CLUSTER

C) El usuario en cuestión tiene el privilegio de sistema UNLIMITED TABLESPACE;

D) El usuario en cuestión tiene el priviliegio de sistema CREATE ANY INDEX .

10) Se puede eliminar un cluster sin eliminar las tablas?

A) SI

B) NO