Universidad de los Andes

Sistemas Transaccionales

Valor Andes

María Paula Mancipe Díaz

Gustavo Adolfo Alegría Zúñiga

Noviembre 09, 2011

**Análisis de Índices**

**Requerimientos y Sentencias SQL**

|  |  |
| --- | --- |
| **REQUERIMIENTO** | **SENTENCIA SQL** |
| CONSULTAR MOVIMIENTOS DE VALORES 1 | SELECT \*  FROM ((((OPERACION\_BURSATIL JOIN  ((INSTRUMENTO\_FINANCIERO JOIN TIPO\_VALOR ON TIPO\_VALOR = TIPO\_VALOR.ID ) JOIN RENTABILIDAD ON ID\_RENTABILIDAD = RENTABILIDAD.ID ) ON ID\_INS\_FIN = INSTRUMENTO\_FINANCIERO.ID) JOIN COMISIONISTA ON COMISIONISTA.NUM\_REGISTRO = ID\_COMISIONISTA\_2)  JOIN INVERSIONISTA ON INVERSIONISTA.ID = ID\_INVERSIONISTA) JOIN OFERENTE ON OFERENTE.ID = ID\_OFERENTE) WHERE FECHA\_FINAL BETWEEN TO\_DATE(‘fechai’) AND TO\_DATE('fechaf) + CRITERIO; |
| CONSULTAR MOVIMIENTOS DE VALORES 2 | SELECT \*  FROM ((((OPERACION\_BURSATIL JOIN  ((INSTRUMENTO\_FINANCIERO JOIN TIPO\_VALOR ON TIPO\_VALOR = TIPO\_VALOR.ID ) JOIN RENTABILIDAD ON ID\_RENTABILIDAD = RENTABILIDAD.ID ) ON ID\_INS\_FIN = INSTRUMENTO\_FINANCIERO.ID) JOIN COMISIONISTA ON COMISIONISTA.NUM\_REGISTRO = ID\_COMISIONISTA\_2)  JOIN INVERSIONISTA ON INVERSIONISTA.ID = ID\_INVERSIONISTA) JOIN OFERENTE ON OFERENTE.ID = ID\_OFERENTE) WHERE FECHA\_FINAL BETWEEN TO\_DATE(‘fechai’) AND TO\_DATE('fechaf) + !CRITERIO; |
| CONSULTAR PORTAFOLIOS | SELECT \*  FROM (PORTAFOLIO  JOIN COMPOSICION ON ID\_PORTAFOLIO = PORTAFOLIO.ID) JOIN  (INSTRUMENTO\_FINANCIERO JOIN TIPO\_VALOR ON TIPO\_VALOR = TIPO\_VALOR.ID )  ON INSTRUMENTO\_FINANCIERO.ID = ID\_VALOR  WHERE TIPO\_VALOR.NOMBRE = ‘&nombre’ AND VALOR > &valor; |
| CONSULTAR VALORES 2 | SELECT \* FROM (PORTAFOLIO  JOIN COMPOSICION ON ID\_PORTAFOLIO = PORTAFOLIO.ID) JOIN  (INSTRUMENTO\_FINANCIERO JOIN TIPO\_VALOR ON TIPO\_VALOR = TIPO\_VALOR.ID )  ON INSTRUMENTO\_FINANCIERO.ID = ID\_VALOR  WHERE  INSTRUMENTO\_FINANCIERO.ID = &aleatorio; |

**Índices, Cantidad de Bloques y Justificación**

|  |  |  |
| --- | --- | --- |
| **CONSULTAR MOVIMIENTOS DE VALORES 1** | | |
| **INDICE** | **BLOQUES** | **JUSTIFICACIÓN** |
| **ID\_INS**  (id del instrumento financiero en OPERACIÓN\_BURSATIL) | 12 | La creación de 12 índices para esta consulta cuenta con varios argumentos:   * El primero de ellos es que consultar movimientos en una bolsa de valores es un requerimiento de alta concurrencia, puesto que los usuarios quieren consultar constantemente las operaciones bursátiles que se han realizado en diferentes horas del día, antes del cierre de la bolsa. Por tanto sería ineficiente para las bases de datos tener que buscar sobre todos los registros que hay, sin que por lo menos hayan estructuras que ya tengan guardados los parámetros. * Un segundo argumento se basa en que los usuarios, ya sean intermediarios, oferentes o inversionistas, esperan una alta selectividad en el momento de buscar las operaciones sobre uno o varios instrumentos financieros. De modo que si se tienen índices dispersos como los que se crearon, la consulta será más selectividad, tomara menos tiempo, y tendrá menor cantidad de operaciones. * La cantidad de espacio que se utiliza para almacenar los índices es una ganancia mientras que se analice, valide y se reconstruyan los índices constantemente de manera que no pierdan la eficiencia. Además la cantidad de espacio es justificada si se consultan operaciones constantemente. De lo contrario, se podrían eliminar índices. * El índice de la fecha tiene una justificación especial. La mayoría de usuarios quieren estar actualizados y saber lo que ha ocurrido en los últimos días o incluso en las últimas horas. Este índice disminuye significativamente el costo mientras que el intervalo de tiempo sea pequeño como el de los usuarios. |
| **FECHA**  (fecha final en OPERACIÓN\_BURSATIL) | 16 |
| **ID\_COMISIONISTA\_2**  (El intermediario del usuario que cierra la operación bursátil en OPERACIÓN\_BURSATIL) | 13 |
| **ID\_OFERENTE**  (El oferente que vende el instrumento financiero en OPERACIÓN\_BURSATIL) | 13 |
| **ID\_INVERSIONISTA**  (El inversionista que compra el instrumento financiero en OPERACIÓN\_BURSATIL) | 13 |
| **NOMBRE\_INSTRUMENTO**  (El nombre del instrumento financiero en INSTRUMENTO\_FINANCIERO) | 80 |
| **RENTABILIDAD\_1**  (El id de la rentabilidad en INSTRUMENTO\_FINANCIERO) | 1 |
| **NOMBRE\_TIPO**  (El nombre del tipo del instrumento financiero en TIPO\_VALOR) | 1 |
| **RENTABILIDAD**  (El nombre del tipo de rentabilidad en RENTABILIDAD) | 1 |
| **COMISIONISTA**  (El nombre del representante en COMISIONISTA) | 139 |
| **INVERSIONISTA**  (El nombre del representante en INVERSIONISTA) | 346 |
| **OFERENTE**  (El nombre del representante en OFERENTE) | 256 |

|  |  |  |
| --- | --- | --- |
| **CONSULTAR MOVIMIENTOS DE VALORES 2** | | |
| **ÍNDICE** | **BLOQUES** | **JUSTIFICACIÓN** |
| Se desactivan 6 índices del requerimiento anterior:  **NOMBRE\_INSTRUMENTO**  **NOMBRE\_TIPO**  **RENTABILIDAD**  **COMISIONISTA**  **INVERSIONISTA**  **OFERENTE** | 0 | Para esta consulta se desactivan seis índices utilizados en el anterior requerimiento:   * Primero porque como se especifica son todos aquellos que NO corresponden al criterio de búsqueda, por tanto si estuviesen activos esto índices haría ineficiente la consulta * Segundo porque los índices no funcionan de manera correcta y no optimizan las búsquedas cuando se utiliza expresiones como “!=” o “NOT IN” que en este caso particular son las que proporcionan respuesta para el requerimiento. Estas expresiones disminuyen la selectividad. |
| **ID\_INS**  (id del instrumento financiero en OPERACIÓN\_BURSATIL) | 12 | Del mismo modo que en el requerimiento anterior, estos índices dispersos sirven para disminuir la cantidad de operaciones, aumentar la selectividad y manejar óptimamente la concurrencia. |
| **FECHA**  (fecha final en OPERACIÓN\_BURSATIL) | 16 |
| **ID\_COMISIONISTA\_2**  (El intermediario del usuario que cierra la operación bursátil en OPERACIÓN\_BURSATIL) | 12 |
| **ID\_OFERENTE**  (El oferente que vende el instrumento financiero en OPERACIÓN\_BURSATIL) | 13 |
| **ID\_INVERSIONISTA**  (El inversionista que compra el instrumento financiero en OPERACIÓN\_BURSATIL) | 13 |
| **RENTABILIDAD\_1**  (El id de la rentabilidad en INSTRUMENTO\_FINANCIERO) | 1 |

|  |  |  |
| --- | --- | --- |
| **CONSULTAR PORTAFOLIOS** | | |
| **INDICE** | **BLOQUES** | **JUSTIFICACIÓN** |
| **ID\_PORTAFOLIO**  (El id del portafolio en COMPOSICION) | 0 | La combinación del id del portafolio y el id del valor es la llave primaria de COMPOSICION y por tanto es el índice creado por defecto en Oracle.  En cuanto al índice disperso sobre el tipo de valor, es un índice que ayudaría a aumentar la selectividad y a filtrar los resultados fácilmente. |
| **TIPO\_VAL**  (El tipo del valor en INSTRUMENTO\_FINANCIERO) | 58 |
| **ID\_VALOR**  (El id del valor en COMPOSICION) | 0 |

|  |  |  |
| --- | --- | --- |
| **CONSULTAR VALORES 2** | | |
| **INDICE** | **BLOQUES** | **JUSTIFICACIÓN** |
| **ID\_PORTAFOLIO**  (El id del portafolio en COMPOSICION) | 0 | La combinación del id del portafolio y el id del valor es la llave primaria de COMPOSICION, el id del instrumento financiero es la llave primaria de INSTRUMENTO FINANCIERO y por tanto son índices creados por defecto en Oracle.  En cuanto al índice disperso sobre el tipo de valor, es un índice que ayudaría a aumentar la selectividad y a filtrar los resultados fácilmente. |
| **TIPO\_VAL**  (El tipo del valor en INSTRUMENTO\_FINANCIERO) | 58 |
| **ID\_VALOR**  (El id del valor en COMPOSICION) | 0 |
| **ID\_INSTRUMENTO\_FINANCIERO**  (El id del instrumento financiero en INSTRUMENTO\_FINANCIERO) | 0 |

**Índices creados por Oracle, Beneficio y Justificación**



Los índices establecidos por Oracle en las tablas PORTAFOLIO, OPERACIÓN\_BURSATIL e INSTRUMENTO FINANCIERO que corresponden a las llaves primarias, son creados para facilitar el manejo y la organización física de la base de datos. Es decir con este índice intentan crear bloques de información en los discos, donde los registros estén ordenados por su llave primaria y no haya segmentos particionados. Sin embargo, a medida que se modifican las tablas y se eliminan los registros, este índice pierde eficiencia. Para esto, constantemente Oracle analiza y valida las estructuras de los índices para aumentar su eficiencia.

Los índices creados por Oracle no son de ayuda para los requerimientos, CONSULTAR MOVIMIENTOS DE VALORES 1 Y 2 porque estos requieren de índices secundarios combinados sobre los parámetros del usuario que no son llaves primarias de las tablas o sencillamente porque no requieren de índices. En los requerimientos de CONSULTAR VALORES 2 y CONSULTAR PORTAFOLIO, si tiene utilidad puesto que entre los parámetros están el identificador del valor que en este caso es llave primaria de INSTRUMENTO\_FINANCIERO y el identificador del portafolio y del valor que combinados actúan como la llave primaria de COMPOSICION y actuarían como un índice en una tabla de Hash. No obstante si no se utiliza otro índice sobre la tabla PORTAFOLIO e INSTRUMENTO FINANCIERO, probablemente la consulta no será tan eficiente como se quiere.

**Requerimientos, Costos y Tiempos**

|  |  |  |
| --- | --- | --- |
| **REQUERIMIENTO** | **COSTO CON ÍNDICES** | **TIEMPO CON INDICES(s)** |
| CONSULTAR MOVIMIENTOS DE VALORES 1  (Por nombre del valor) | 8 | 0,599 |
| CONSULTAR MOVIMIENTOS DE VALORES 1  (Por tipo del valor) | 18 | 0,39 |
| CONSULTAR MOVIMIENTOS DE VALORES 1  (Por tipo de rentabilidad) | 23 | 0,508 |
| CONSULTAR MOVIMIENTOS DE VALORES 1  (Por nombre del comisionista) | 8 | 0,422 |
| CONSULTAR MOVIMIENTOS DE VALORES 1  (Por nombre del oferente) | 305 | 0,39 |
| CONSULTAR MOVIMIENTOS DE VALORES 1  (Por nombre del inversionista) | 8 | 0,444 |
| CONSULTAR MOVIMIENTOS DE VALORES 2  (Por nombre del valor) | 5818 | 0,454 |
| CONSULTAR MOVIMIENTOS DE VALORES 2  (Por tipo del valor) | 5818 | 0,383 |
| CONSULTAR MOVIMIENTOS DE VALORES 2  (Por tipo de rentabilidad) | 5818 | 0,401 |
| CONSULTAR MOVIMIENTOS DE VALORES 2  (Por nombre del comisionista) | 5818 | 0,384 |
| CONSULTAR MOVIMIENTOS DE VALORES 2  (Por nombre del oferente) | 5818 | 0,372 |
| CONSULTAR MOVIMIENTOS DE VALORES 2  (Por nombre del inversionista) | 5818 | 0,452 |
| CONSULTAR PORTAFOLIOS | 24 | 0,536 |
| CONSULTAR VALORES 2 | 4 | 0,528 |

**Distribución de parámetros**

La distribución de los parámetros afecta en particular a los requerimientos 1 y 2 puesto que estas contienen fechas que son elegidas por el usuario de Valor Andes. Cuando se ejecuta el plan en Oracle con dos parámetros aleatorios de fechas no es posible visualizar ningún cambio. Sin embargo, cuando se ejecutan los casos manualmente se puede ser que cuando el intervalo entre la fecha inicial y la final es amplio, aumenta el tiempo de la consulta, por el contrario cuando es corto, se disminuye significativamente el tiempo. En el resto de los parámetros la distribución no afecta la manera en que se comportan las consultas puesto que hay la misma probabilidad de que el nombre del oferente en el requerimiento 1 comience por A o por Z.

**Análisis de consultas con parámetros aleatorios**

Para realizar el análisis de las consultas, se tomaron valores aleatorios en los planes de ejecución como si el usuario fuera el que estuviera realizando la consulta en SQL developper. Esto se decidió basado en que solo los parámetros de fechas, afectan el rango del resultado, es decir si el intervalo es mayor, la búsqueda es más compleja. El resto de los parámetros pueden ser totalmente aleatorios y no hay ninguna preferencia. Un usuario podría consultar desde las operaciones de una acción de Davivienda, hasta las de un bono del Banco de la Republica.

En todo caso, así fuesen unos parámetros específicos, la diferencia de tiempo no es significativa incluso cuando no se utilizan índices como en el requerimiento 2. Como se puede ver la diferencia es 0,005 aproximadamente, un tiempo que no representa una diferencia significativa puesto que el tiempo de respuesta de un ser humano es entre 2 y 3 segundos. De modo que para la consulta, y compra de valores el tiempo de las consultas es óptimo.

**Escenarios de Prueba**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CONSULTAR MOVIMIENTOS DE VALORES 1** | | | | | |
| Nombre del Parámetro | Valor 1 | Valor 2 | Tiempo 1 | Tiempo 2 | Tiempo del Plan |
| Fecha Inicial | 09/11/1980 | 09/11/1981 | 0,187 | 0,172 | 0,468 |
| Fecha Final | 10/11/2014 | 10/11/2012 | 0,062 | 0,171 | 0,406 |
| Nombre del valor | Zoonder | Rooxo | 0,172 | 0,174 | 0,599 |
| Tipo del valor | Bono Publico | Acción | 0,140 | 0,163 | 0,390 |
| Tipo de rentabilidad | Rentabilidadii | Rentabilidadfi | 0,124 | 0,156 | 0,508 |
| Nombre del comisionista | David Ruiz | Elizabeth Ch | 0,179 | 0,193 | 0,422 |
| Nombre del oferente | Joyce Hansen | Nicole Bennet | 0,125 | 0,203 | 0,390 |
| Nombre del inversionista | Jerry Lane | Brian Lee | 0,201 | 0,194 | 0,444 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CONSULTAR MOVIMIENTOS DE VALORES 2** | | | | | |
| Nombre del Parámetro | Valor 1 | Valor 2 | Tiempo 1 | Tiempo 2 | Tiempo del Plan |
| Fecha Inicial | 09/11/1980 | 09/11/1981 | 0,187 | 0,172 | 0,468 |
| Fecha Final | 10/11/2014 | 10/11/2013 | 0,062 | 0,171 | 0,406 |
| Nombre del valor | Zoonder | Rooxo | 0,203 | 0,205 | 0,454 |
| Tipo del valor | Bono Público | Acción | 0,187 | 0,190 | 0,383 |
| Tipo de rentabilidad | Rentabilidadii | Rentabilidadfi | 0,156 | 0,140 | 0,401 |
| Nombre del comisionista | David Ruiz | Elizabeth Ch | 0,187 | 0,188 | 0,384 |
| Nombre del oferente | Joyce Hansen | Nicole Bennet | 0,390 | 0,187 | 0,372 |
| Nombre del inversionista | Jerry Lane | Brian Lee | 0,312 | 0,287 | 0,452 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CONSULTAR PORTAFOLIO** | | | | | |
| Nombre del Parámetro | Valor 1 | Valor 2 | Tiempo 1 | Tiempo 2 | Tiempo del Plan |
| Tipo de valor | Bono Publico | CDT | 0,094 | 0,094 | 0,536 |
| Valor | 7000 | 10000 | 0,093 | 0,098 | 0,536 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CONSULTAR VALORES** | | | | | |
| Nombre del Parámetro | Valor 1 | Valor 2 | Tiempo 1 | Tiempo 2 | Tiempo del Plan |
| Id del valor | 78 | 20000 | 0,094 | 0,078 | 0,528 |

Con los casos de prueba que se presentaron anteriormente se pueden observar varios aspectos:

1. El tiempo que se gasta en ejecutar/calcular el plan Oracle, es el tiempo máximo que se gastaría en el peor de los casos. Con cada uno de los valores que fueron asignados en los casos de prueba se gasta la mitad o incluso menos del valor estimado por el plan.
2. Cuando los índices son combinados, secundarios y dispersos la consulta toma mucho más tiempo para ubicar los resultados. Esto se debe a que los índices secundarios se almacenan en árboles B+, esto implica que a pesar de su eficiencia, hay que recorrerlos. Este recorrido gasta aproximadamente log(n), lo que hace que la consulta dure un poco más.
3. Los índices primarios que son creados por Oracle y que son llaves primarias en las tablas utilizados en consultas gastan la mitad de tiempo o menos que las consultas con índices secundarios. Esto radica en que los índices primarios se basan en tablas de hashing, así que no es necesario recorrer una estructura, sino que se tiene el índice directo para llegar al registro.
4. La diferencia de tiempos entre dos valores de un mismo parámetro es demasiado pequeña puesto que como se mencionaba antes hay la misma probabilidad de que un parámetro comience por A que por Z o que sea 0 a que 7688. De modo usar un valor aleatorio en los planes de ejecución nos ayuda a saber cuánto es el máximo de tiempo que nos podemos gastar.