**Parcial sobre sistemas de archivos**

**Integrantes : Jaime David Mejia Quintero**

**Eduar Cortes Tamayo**

**Sebastian Eusse Bermudez**

**Docente : Sergio Armando Gutierrez Betancur**

**Universidad Nacional de Colombia**

**Sede Medellin**

**2015 2**

Experimento.

Dada la maquina virtual con la distribucion de Unix Debian, se hicieron diferentes controladores cada uno con disco respectivo. En general la distribucion que se programo para cada de los controladores y su diferentes discos son los siguientes :

sda disco 10

sdb disco 11

sdc disco 12

sdd disco 13

sde disco 20

sdf disco 21

sdg disco 22

sdh disco 23

sdi disco principal del sistema

sdj copia del disco principal

md0 raid0

md1 raid 0 strriping

md2 raid 1

Los discos con la primera controladora se encuentra en el formato SCSI, los discos con la segunda controladora se encuentran en formato SAS, el disco principal y la copia de este se encuentran bajo el formato SATA.

Luego de la instalacion de las controladoras y los discos se procedio a instalar los paquetes y a denotar con nombre cada uno de los discos(explicado en la parte de arriba). Se realizaron las particiones que fueron pedidas en el experimento (disco x0 y x2 con una única partición, disco x1 y x3 con dos particiones de 250 MB).

Fueron configurados los discos para que fueran de tipo RAID(raid0,raid1,raid0 con Stripping). La primera diferencia es ver como el raid1 se acortaba de tamaño. Esto se debe a que raid1 es una configuración de seguridad el cual guardaba la mitad del espacio del disco duro para realizar copias de seguridad del sistema. En el caso de raid0 y raid0 con stripping no hubo mucha diferencia con los cambios de tamaño del disco duro.

Ahora con la creación de los discos RAID falta montar los diferentes sistemas de archivos para la prueba, el primero sera el EXT4. Se realiza la configuración para los permisos necesarios para mysql e igualmente para un usuario por medio de un enlace simbólico.

Se preparo la base de datos y finalmente se empieza a hacer el monitoreo de la base de datos con el comando de vmstat y iostat, Para cada comando se utilizaron las siguientes propiedades :

**vmstat 2 > vmstat.txt**

**iostat 2 -> iostat.txt**

Cada 2 segundos guardaba la información en un txt, la información queda almacenada de la siguiente manera (el primero para el iostat,el segundo para el vmstat):

**Con iostat(para el ext4) :**

avg-cpu: %user %nice %system %iowait %steal %idle

11,00 0,01 1,80 1,78 0,00 85,41

Device: tps kB\_read/s kB\_wrtn/s kB\_read kB\_wrtn

sda 0,36 2,15 18,31 5979 50900

**sdb 0,96 94,51 15,23 262745 42336**

sdc 0,34 2,18 8,98 6051 24968

sdd 0,19 1,46 0,00 4048 4

sde 0,21 1,46 0,04 4071 120

sdf 0,76 1,54 106,88 4288 297116

sdg 0,44 3,05 8,82 8466 24508

sdh 0,19 1,46 0,00 4048 4

sdi 4,36 119,13 5,21 331169 14488

sdj 0,11 0,85 0,00 2361 8

md0 1,78 0,63 23,38 1739 65004

md1 1,80 0,63 23,35 1750 64904

md2 0,73 0,73 15,22 2041 42304

**Con iostat(para el ext3) :**

Device: tps kB\_read/s kB\_wrtn/s kB\_read kB\_wrtn

sda 0,26 1,56 13,30 5979 50900

**sdb 0,70 68,67 11,07 262746 42336**

sdc 0,25 1,59 6,54 6083 25012

sdd 0,13 1,06 0,00 4048 4

sde 0,15 1,06 0,03 4071 120

sdf 0,55 1,12 77,66 4288 297116

sdg 0,32 2,21 6,41 8474 24512

sdh 0,13 1,06 0,00 4048 4

sdi 23,37 157,91 558,23 604165 2135796

sdj 0,08 0,62 0,00 2361 8

md0 1,29 0,45 16,99 1739 65004

md1 1,31 0,47 16,98 1790 64952

md2 0,53 0,53 11,06 2042 4230

**Con iostat(para el xfs) :**

Device: tps kB\_read/s kB\_wrtn/s kB\_read kB\_wrtn

sda 0,26 1,55 14,42 8170 75998

**sdb 0,59 50,06 11,84 263853 62412**

sdc 0,23 1,25 8,51 6585 44863

sdd 0,10 0,77 0,00 4048 4

sde 0,16 0,86 3,31 4511 17459

sdf 0,44 0,82 60,18 4328 317192

sdg 0,27 1,62 8,22 8531 43338

sdh 0,10 0,77 0,00 4048 4

sdi 38,64 169,42 734,35 893021 3870848

sdj 0,06 0,45 0,00 2361 8

scd0 0,01 0,19 0,00 978 0

md0 3,05 0,83 20,38 4370 107441

md1 2,58 0,45 19,66 2349 103629

md2 1,06 0,60 11,83 3189 62375

**Con vmstat (ext4):**

r b swpd free buff cache si so bi bo in cs us sy id wa st

0 0 96716 55588 66988 137564 1 26 170 596 127 892 14 3 77 6 0

0 0 96716 55588 66988 137584 0 0 8 0 142 466 19 3 77 1 0

1 0 96716 54952 66992 138116 4 0 264 0 262 835 41 6 49 5 0

2 0 96688 52836 67000 140104 24 0 1054 16 273 526 40 5 44 11 0

1 0 96688 52564 67000 140104 0 0 0 0 236 545 41 2 57 0 0

2 0 96688 52424 67000 140104 0 0 10 0 76 345 11 2 87 1 0

2 0 96688 52332 67008 140120 0 0 0 28 297 793 54 5 41 0 0

2 0 96688 52208 67008 140128 0 0 6 0 375 1145 74 6 20 0 0

0 0 96688 50464 67008 141828 4 0 848 0 256 734 41 4 45 10 0

1 0 97932 60556 66576 133056 2 622 34 650 179 672 37 5 56 3 0

0 0 97932 60564 66576 133032 0 0 0 0 92 304 15 2 83 0 0

0 0 97928 60440 66576 133136 2 0 2 0 236 973 36 6 57 1 0

1 0 97928 60420 66584 133080 0 0 0 8 116 420 10 3 86 2 0

0 0 97928 60420 66584 133088 0 0 0 0 99 362 12 2 86 0 0

0 0 97928 60544 66584 133036 0 0 0 0 86 335 11 2 87 0 0

0 0 97928 60576 66592 133136 0 0 64 8 82 381 13 1 85 1 0

0 0 97928 60576 66592 133164 0 0 0 2 62 275 6 1 93 0 0

0 0 97928 60568 66592 133168 0 0 0 0 70 259 8 1 91 0 0

3 0 97928 60576 66600 133160 0 0 0 30 68 274 6 1 93 0 0

1 0 97928 60576 66600 133232 0 0 0 0 296 1046 35 7 58 0 0

**Con vmstat (ext3):**

r b swpd free buff cache si so bi bo in cs us sy id wa st

0 0 121664 105272 52920 126192 2 29 160 683 144 1112 16 4 74 6 0

1 0 121664 105232 52920 126140 0 0 0 0 167 499 26 2 72 0 0

0 0 121664 105216 52920 126140 0 0 0 2 130 457 22 2 76 0 0

1 0 121664 104984 52928 126204 0 0 34 26 131 509 17 2 75 6 0

0 0 121664 104984 52928 126212 0 0 0 0 198 678 39 4 57 0 0

0 0 121664 104868 53060 126260 0 0 88 0 115 484 16 2 79 3 0

0 0 121664 104868 53068 126256 0 0 0 26 92 295 12 2 84 3 0

0 0 121664 104868 53068 126264 0 0 0 0 74 274 11 1 88 0 0

5 0 121664 102016 53904 127136 0 0 438 1098 609 7201 52 15 31 2 0

6 1 121664 99636 55824 127512 0 0 206 2498 1041 14762 67 30 0 3 0

1 1 121664 98224 57032 127784 0 0 122 1710 774 9391 60 26 0 14 0

7 2 121664 96600 58436 127936 0 0 52 2402 915 11869 79 21 0 0 0

4 1 121664 94828 59936 128060 0 0 40 2458 941 11987 73 27 0 1 0

5 1 121664 93400 61288 128080 0 0 0 2474 881 11125 73 26 0 1 0

7 1 121664 92044 62632 128092 0 0 16 2534 805 11984 70 20 0 10 0

0 1 121664 91136 63584 128156 0 0 16 2116 656 9098 67 17 0 16 0

1 1 121664 90104 64488 128268 0 0 8 2106 657 8110 65 18 0 16 0

1 2 121664 89100 65416 128376 0 0 16 2312 640 8561 57 19 0 24 0

7 0 121664 87952 66312 128456 0 0 44 2426 655 7610 66 18 0 16 0

1 1 121664 86756 67512 128468 0 0 18 2382 869 10263 77 23 0 0 0

5 0 121664 85052 68960 128700 0 0 16 3200 893 12497 65 26 0 8 0

**Con vmstat (xfs):**

r b swpd free buff cache si so bi bo in cs us sy id wa st

1 0 143200 72416 21500 201544 3 28 180 782 171 1440 17 4 72 7 0

0 0 143200 72200 21500 201580 0 0 8 6 117 548 16 1 82 1 0

0 0 143200 72052 21504 201652 14 0 52 0 346 1323 59 5 34 3 0

0 0 143200 71936 21512 201700 0 0 26 26 261 991 52 4 41 3 0

0 0 143200 71936 21512 201708 0 0 0 0 89 420 13 3 84 0 0

0 0 143200 71936 21512 201708 0 0 0 0 70 373 11 2 87 0 0

0 0 143200 71936 21520 201700 0 0 0 34 87 405 14 2 84 1 0

1 0 143200 71764 21524 201712 0 0 12 0 54 289 7 1 91 1 0

0 0 143200 71664 21524 201756 0 0 20 0 367 1318 64 6 27 3 0

0 0 143200 71540 21532 201840 0 0 42 24 187 928 32 3 64 1 0

0 0 143200 70176 21532 203172 0 0 662 2 298 1328 62 5 23 10 0

1 0 143200 69188 21532 203848 0 0 340 2 286 639 44 2 42 12 0

0 0 143200 69188 21540 203848 0 0 0 18 52 147 2 0 97 1 0

0 0 143200 69188 21540 203852 0 0 0 0 74 179 2 1 97 0 0

0 0 143200 68940 21540 203876 0 0 12 0 182 516 34 2 63 2 0

0 0 143200 68940 21552 203912 0 0 20 8 130 634 20 3 71 6 0

0 0 143200 68816 21552 203912 0 0 0 2 82 342 10 2 88 0 0

0 0 143200 68816 21552 203916 0 0 0 0 58 307 6 1 94 0 0

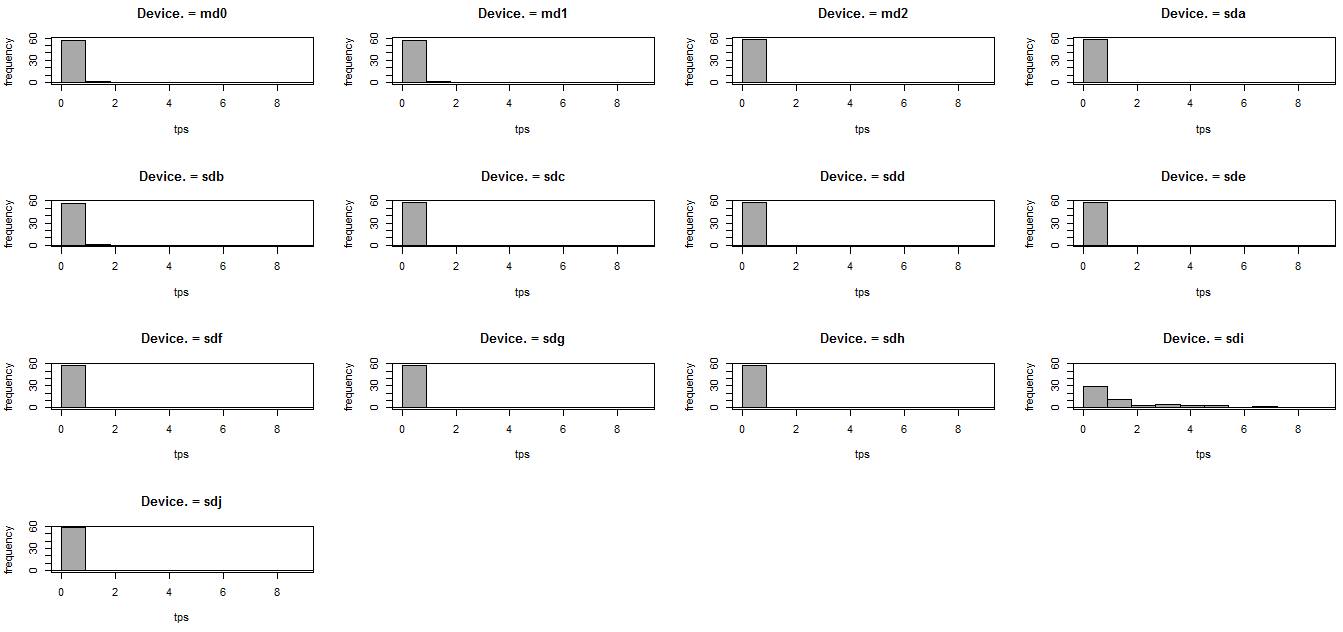
0 0 143200 68816 21560 203916 0 0 0 78 112 515 18 3 75 4 0

3 1 143200 68816 21560 203916 0 0 4 0 51 286 8 1 91 0 0

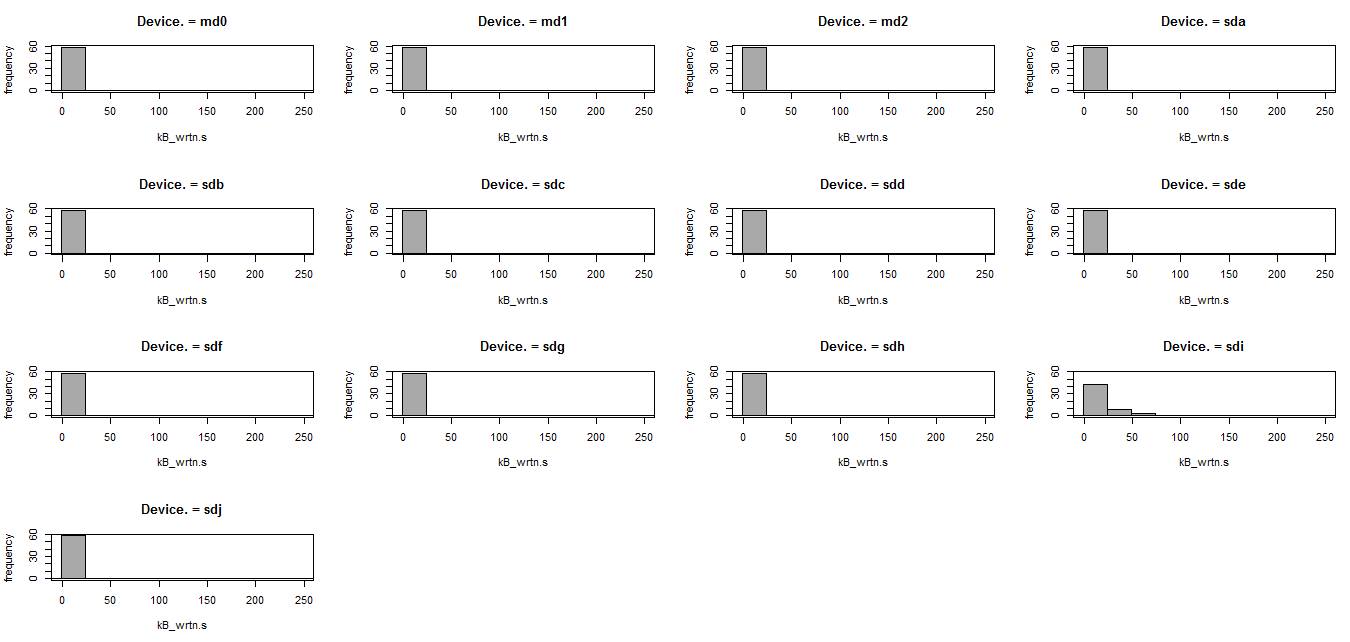
2 1 143200 65120 22812 205312 0 0 620 1450 813 9780 62 21 3 14 0

Para el calculo de los sistemas de archivos ext4,ext3,xfs en el iostat se compara cada uno de los campos presentados arriba, aunque principalmente solo mostraremos tres datos(tps,kb\_wrtn/s,kb\_read/s).

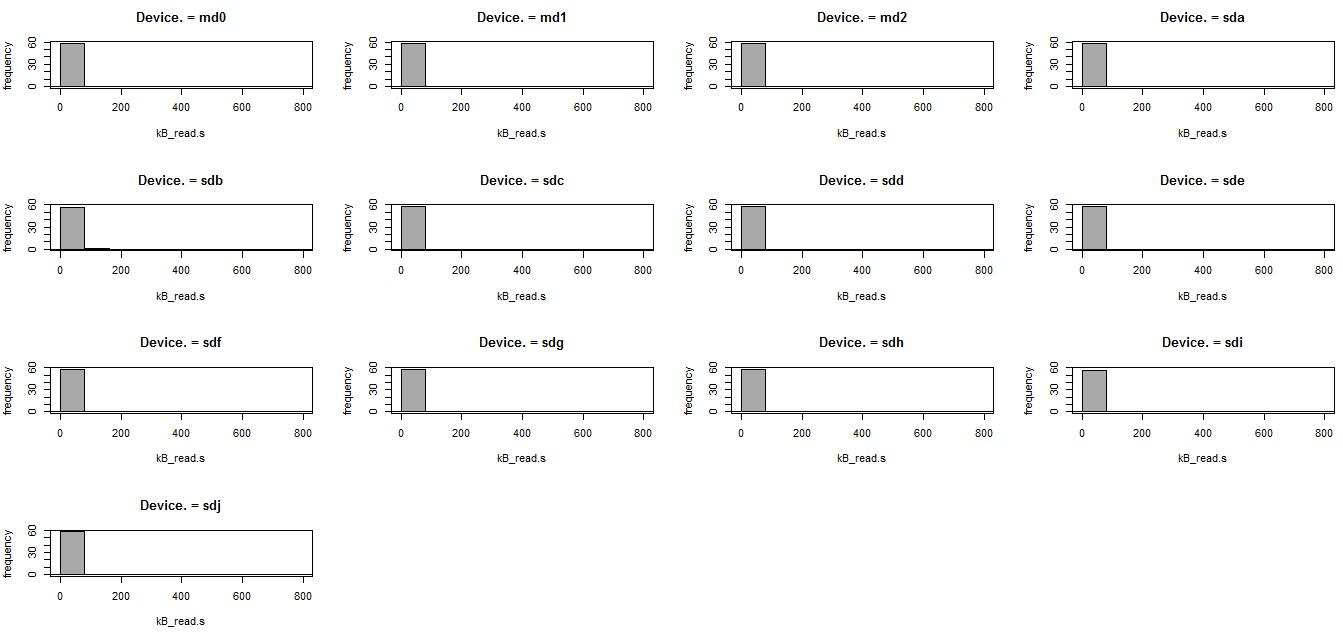
**Datos del tps**



**Datos del KB\_Wrtn/s**



**Datos del KB\_Read/s**



El mejor de todos en el iostat se presento con el sdb, es decir el disco 11 con el formato RAID1. Puede presentarse debido a que se dividió en dos disco con un tamaño pequeño, la especialidad de los RAID1 es poder utilizar el primer disco para la información y el segundo para una copia de los datos, ya que ambos estaban divididos no se gasto tiempo adicional. Igualmente para el vmstat evaluando la parte del disco, el menor gasto se presenta para el EXT4,podemos decir que las diferentes pruebas son parecidas, el formato EXT4 presenta mejores resultados que el EXT3 y el XFS.

Finalmente se desmontaron los dispositivos Raid de las pruebas anteriores y nuevamente colocados pero con otras especificaciones RAID0 con stripping entre los discos disco11 y disco13. Cree un dispositivo RAID1 entre los discos disco21 y disco23 (con formato ext4), se monitores la información con iostat y vmstat, la información otorgada fue la siguiente :

**Con iostat (ext4):**

Device: tps kB\_read/s kB\_wrtn/s kB\_read kB\_wrtn

sdb 0,42 2,68 8,26 3265 10060

sdc 0,12 0,78 0,00 954 0

sda 0,26 1,86 0,00 2266 0

sdd 0,37 2,40 8,07 2918 9824

sde 0,26 1,66 0,00 2016 0

**sdf 3,03 423,34 20,22 515666 24625**

sdg 0,12 0,59 0,00 713 0

sdh 2,89 2,90 440,29 3532 536305

sdi 10,26 269,82 7,91 328660 9632

sdj 0,24 1,56 0,00 1895 0

md0 3,68 0,71 28,58 861 34816

md1 2,21 1,31 20,20 1598 24608

**Con vmstat(ext4):**

r b swpd free buff cache si so bi bo in cs us sy id wa st

1 0 1156 61476 25420 232628 0 1 283 43 93 317 9 2 84 4 0

0 0 1156 61352 25428 232620 0 0 0 14 140 379 15 3 81 1 0

0 2 1156 54820 25476 235728 0 0 1562 64 278 1311 46 5 31 17 0

2 2 1380 57984 25804 229228 0 112 2042 474 310 2668 21 6 0 73 0

1 3 1380 51164 26044 232864 0 0 1814 266 269 2010 17 5 0 77 0

0 3 4928 61464 26204 217248 0 1774 1566 1980 251 1654 16 9 0 75 0

0 4 4928 54588 26508 220828 0 0 1794 422 295 2372 19 6 0 76 0

1 3 4928 71800 16164 214140 0 0 1808 334 298 2182 20 7 0 73 0

0 2 4928 65308 16436 217572 0 0 1684 706 274 2392 12 5 0 82 0

0 2 4928 59620 16696 220740 0 0 1572 458 241 2130 9 5 0 86 0

0 3 4928 52652 17008 224280 0 0 1782 436 279 2549 11 5 0 84 0

0 2 4928 61368 12096 217768 0 0 1542 422 265 2501 10 6 0 84 0

0 3 4928 53652 12472 222216 0 0 2202 536 304 2802 12 7 0 81 0

0 1 4928 60560 8200 216980 0 0 1588 738 267 2462 14 4 1 81 0

2 2 4928 54724 8496 219748 0 0 1388 532 256 2412 13 3 0 84 0

1 1 4928 62752 6648 211144 0 0 1568 502 265 2522 13 4 0 83 0

0 2 4928 55308 7048 215004 0 0 1876 648 334 3641 18 7 0 75 0

2 3 4928 63356 5640 206028 0 0 1352 690 325 3735 16 8 0 76 0

4 3 4928 56264 6136 209484 0 0 1530 1150 367 4349 20 10 0 70 0

4 0 4928 51404 6544 211812 0 0 1124 668 287 3443 14 8 0 78 0

0 1 4928 59732 6364 201648 0 0 1066 770 281 3480 16 6 0 78 0

Esta vez, el disco21 fue mejor que el resto, al igual que en el anterior el disco21 también posee un RAID1 con EXT4. Podemos sacar como conclusión que el formato RAID1 opero mucho mejor que el RAID0 y el RAID0 con striping, además el sistema de archivos EXT4 es mucho mejor que EXT3 y XFS.

**Recomendaciones**

Por ahora, ext4 parece ser la mejor opción en un sistema de escritorio, puesto que ya está presente como un sistema de archivos por defecto, y es ligeramente más rápido que los otros dos al transferir archivos. Pero tenga en cuenta que se trata de una sola unidad en un sistema de escritorio.

Ademas no se podría dar un resultado concreto sobre cual es mejor ya que al realizar las pruebas en una virtualizacion de una maquina, esta no siempre presenta una realidad sobre lo que realizan los discos en un sistema físico verdadero pero en este se demuestra que Ext4 fue el sistema de archivo que obtuvo mejor rendimiento en la maquina virtual donde se realizo la prueba y bajo el sistema operativo debían

Se recomienda mirar otros sistemas de archivos, como el reciente BTRFS que según expertos puede tener mejor perfomance. Probar otros RAID como el 2,3,0+1 . Otras configuraciones distintas podrían presentar mejores resultados

**Anexos**

1) El concepto de partición raw (partición aleatoria) consiste en que la sección del disco duro a trabajar (en el área de partición) no posee un sistema de archivos predeterminado. Es decir no tiene que pasar por las dependencias del sistema operativo. Dado que no hay restricciones, la base de datos puede ser manipulada abiertamente. Habrá que Considerar el riesgo de mala manipulación por parte del usuario.

**Ventajas:** No están sujetos a ningún bloqueo del sistema operativo.

El buffer cache del sistema operativo o cache pasa por alto, dando ganancias de rendimiento y consumo de memoria reducido.

Múltiples sistemas pueden ser fácilmente compartidos.

El sistema de aplicación o base de datos tiene el control total de manipular los componentes.

**Desventajas:** La unidad de asignación a la base de datos es la partición cruda completa. Una partición cruda no puede ser utilizada para múltiples espacios de tablas. Una partición cruda no es lo mismo que un sistema de archivos donde los archivos se pueden crear.

Los administradores tienen que crearlos con un tamaño específico. Cuando las bases de datos crecen en tamaño, las particiones crudas no se pueden extender. Las particiones se deben agregar para apoyar los espacios de tablas de crecimiento. A veces puede haber limitaciones en el número total de particiones sin formato que puede ser utilizado en el sistema. Además, no hay operaciones de base de datos que puede ocurrir en un archivo de datos individual. No existe, por tanto, ningún beneficio lógico de tener un tablespace que constan de muchos data file a excepción de los tablespace que son más grandes que el máximo que Oracle puede soportar en un solo archivo.

Los comandos estándar de manipulación de archivos no se pueden utilizar en las particiones crudas y por lo tanto en los data files. Comandos como cpio o tar no puede ser utilizados para realizar copias de seguridad por lo tanto la estrategia de copia de seguridad será más complicada. Las particiones primas no pueden ser utilizadas para escribir los registros de archivo.

2)Hay que tener en cuenta cómo actúa el Raid 0 y el Raid 1, el Raid 0 aprovecha los estados de los disco duro para colocar información equitativamente lo que proporciona mayor rendimiento. En cambio el Raid 1 actúa como un dispositivo de respaldo del disco original (una copia).

Dado estos casos, para el raid 0, la mayor ventaja es que la diferencia entre los archivos se toma por su media (si un disco de 300 GB se divide con uno de 100 GB, el tamaño del conjunto resultante será sólo de 200 GB, ya que cada disco aporta 100GB). El problema es que en caso de daño posiblemente no se pueda recuperar la información.

Para el Raid 1, su principal ventaja es la recuperación de archivos en caso de problemas con el disco. Su desventaja es que la porción de archivo del disco mas pequeño es la que puede ser recuperada (es decir si el disco es de 40GB y el disco mayor es de 160GB solamente 40GB pueden ser recuperados del resto son inutilizados).

3)El principal problema que se puede tener es cuando deseas mover una partición que se encuentra en una sección del disco a otra. Principalmente porque Luego de la partición, cada una es independiente al resto. Sin embargo, si se desea tener un espacio libre tendrá que formatear el espacio del disco duro correspondiente a eliminar, y colocar el formato del disco adecuado (si existe incompatibilidad entre los sistemas de archivo tendrá error).

Además por tener varios sistemas de archivos el arranque será más lento (como cada disco es independiente, se tendrá que hacer una mayor cantidad de comprobaciones).

4) La máscara del dir es la 771, es decir. Los permisos que posee el directorio son: RWX RWX –X (de izquierda a derecha seria el usuario, grupo y otros). Para la máscara del file es decir, la máscara 444 seria: R-- R-- R-- (archivos de solo lectura para el archivo). Teniendo esto en cuenta. Los siguientes comandos para ejecutarse serian

A) ls dir: principalmente el comando ls permite listar las ramas donde se está indicando mediante el nombre y algunos atributos. Como dir posee ejecución para otros usuarios podrá comprobarse normalmente lo que esté pasando el sistema.

B) ls -d dir : Solo muestra el nombre sin necesidad de entrar a ver que tiene en el interior para ver si ese directorio existe. También desde un usuario externo se puede acceder.

C)ls -l dir/file : acceder mediante un formato long a la dirección dir/file. En cuanto a la máscara de direcciones para dir puede acceder debido a que tiene ejecución, para la máscara de direcciones de file no se podrá ver tanto la ejecución del archivo como la modificación u escritura de este.

D)cat dir/file: El comando cat nos permite mirar el archivo, como file posee permisos para lectura el comando podrá visualizar el archivo sin problemas

E)su -c “chown usr3 dir” usr : Se ejecuta desde el administrador principal el cual posee todos los permisos, la opción -c muestra la lista por orden cronológico, finalmente “chown usr3 dir” usr permite cambiar el archivo desde un usuario a otro, como se está ejecutando desde el administrador el cual posee todo los permisos, la ejecución seria que se cambia el directorio desde usr hasta usr3.

5) Se registra que es de 1% debido a que la limpieza del antiguo formato y la creación de uno nuevo. Es decir, el proceso mostrado empieza desde la creación de una partición, el cambio del formato y montar dicho formato. Quedando una partición del disco duro nuevo con un nuevo formato el cual de todo el espacio de memoria que posee solamente se ha utilizado un 1%.