



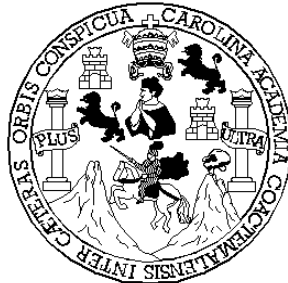
Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

**UTILIZACIÓN DE HERRAMIENTAS DE MONITOREO
PARA LA EVALUACIÓN Y RENDIMIENTO DE SISTEMAS
OPERATIVOS**

Saúl Francisco Choc González
Asesorado por: Ing. Jorge Luis Álvarez

Guatemala, octubre de 2004

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**UTILIZACIÓN DE HERRAMIENTAS DE MONITOREO
PARA LA EVALUACIÓN Y RENDIMIENTO DE SISTEMAS
OPERATIVOS**

TRABAJO DE GRADUACIÓN

PRESENTADO A JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERIA

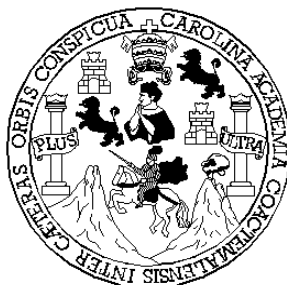
POR

SAÚL FRANCISCO CHOC GONZÁLEZ
ASESORADO POR ING. JORGE LUIS ÁLVAREZ

AL CONFERÍRSELE EL TÍTULO DE
INGENIERO EN CIENCIAS Y SISTEMAS

GUATEMALA, OCTUBRE DE 2004

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Sydney Alexander Samuels Milson
VOCAL I	Ing. Murphy Olympto Paiz Recinos
VOCAL II	Lic. Amahán Sánchez Álvarez
VOCAL III	Ing. Julio David Galicia Celada
VOCAL IV	Br. Kenneth Issur Estrada Ruiz
VOCAL V	Br. Elisa Yazminda Vides Leiva
SECRETARIO	Ing. Pedro Antonio Aguilar Polanco

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Sydney Alexander Samuels Milson
EXAMINADOR	Ing. Otto René Escobar Leiva
EXAMINADOR	Ing. Virginia Victoria Tala Ayerdi
EXAMINADOR	Ing. Jorge Armín Mazariegos Rabanales
SECRETARIO	Ing. Pedro Antonio Aguilar Polanco

HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

**UTILIZACIÓN DE HERRAMIENTAS DE MONITOREO
PARA LA EVALUACIÓN Y RENDIMIENTO DE SISTEMAS OPERATIVOS**

Tema que me fuera asignado por la Dirección de la Escuela de Ingeniería en Ciencias y Sistemas con fecha febrero 2004.

Saúl Francisco Choc González

Guatemala, Septiembre 4 de 2004

Ingeniero
Carlos Azurdia
Coordinador de Privados y Revisión de Tesis
Escuela de Ciencias y Sistemas

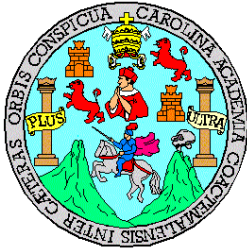
Estimado Ingeniero:

Por medio de la presente, me permito informarle que he asesorado el trabajo de graduación titulado: **UTILIZACIÓN DE HERRAMIENTAS DE MONITOREO PARA LA EVALUACIÓN Y RENDIMIENTO DE SISTEMAS OPERATIVOS**, elaborado por el estudiante Saúl Francisco Choc González, y a mi juicio el mismo cumple con los objetivos propuestos para su desarrollo.

Agradeciéndole de antemano la atención que le preste a la presente, me suscribo de usted.

Atentamente,

Ing. Jorge Luis Álvarez
Asesor



Universidad San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 06 de Septiembre de 2004

Ingeniero
Luis Alberto Vettorazzi España
Director de la Escuela de Ingeniería
En Ciencias y Sistemas

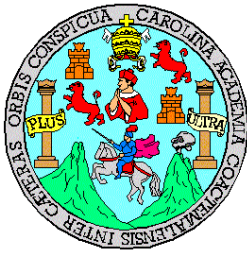
Respetable Ingeniero Vettorazzi:

Por este medio hago de su conocimiento que he revisado el trabajo de graduación del estudiante **SAUL FRANCISCO CHOC GONZALEZ**, titulado: **“UTILIZACIÓN DE HERRAMIENTAS DE MONITOREO PARA LA EVALUACIÓN Y RENDIMIENTO DE SISTEMAS OPERATIVOS”**, y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo.

Al agradecer su atención a la presente, aprovecho la oportunidad para suscribirme,

Atentamente

Ing. Carlos Alfredo Azurdia
Coordinador de Privados
Y Revisión de Trabajos de Graduación



Universidad San Carlos de Guatemala
Facultad de Ingeniería

EL Director de la Escuela de Ingeniería en Ciencias y Sistemas de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor con el visto bueno del revisor y del Licenciado en Letras, del trabajo de graduación titulado **“UTILIZACIÓN DE HERRAMIENTAS DE MONITOREO PARA LA EVALUACIÓN Y RENDIMIENTO DE SISTEMAS OPERATIVOS”**, presentado por el estudiante **Saúl Francisco Choc González**, aprueba el presente trabajo y solicita la autorización del mismo.

ID Y ENSEÑAD A TODOS

Ing. Luis Alberto Vettorazzi España
DIRECTOR
INGENIERIA EN CIENCIAS Y SISTEMAS

Guatemala, 29 de septiembre de 2004

Ref. DTG-350-2004

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería en Ciencias y Sistemas, al Trabajo de Graduación titulado: **UTILIZACIÓN DE HERRAMIENTAS DE MONITOREO PARA LA EVALUACIÓN Y RENDIMIENTO DE SISTEMAS OPERATIVOS**, presentado por el estudiante universitario, **Saúl Francisco Choc González**, procede a la autorización para la impresión del mismo.

IMPRIMASE:

Ing. Sydney Alexander Samuels Milson
DECANO

Guatemala, octubre del 2004

/lmcb.

AGRADECIMIENTOS A

Dios

Por permitirme coronar de esta manera mis estudios.

Ing. Jorge Luis Álvarez

Por su apoyo y ayuda incondicional en el desarrollo de este trabajo de graduación, eternamente agradecido.

Mis amigos

Jonatan, Guaya, Tota, Topo, Chino Lau, Tres, Don D'Wilson, Julito, con quienes hemos pasamos momentos inolvidables.

Mis padres

Por su esfuerzo, y amor en todo momento.

La Universidad de San Carlos de Guatemala

Por la preparación y conocimientos adquiridos por medio de los diferentes ingenieros quienes han realizado una loable labor.

A mí querida esposa Beatriz y a nuestro amado hijo Geralf, con quienes Dios me ha bendecido grandemente, afinando y monitoreando nuestra vida diariamente.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	XI
LISTA DE SÍMBOLOS	XIV
GLOSARIO	XV
RESUMEN	XXII
OBJETIVOS	XXIII
INTRODUCCIÓN	XXV
1 FUNDAMENTOS TEÓRICOS DE MEDICIÓN, EVALUACIÓN, Y HERRAMIENTAS DE MONITOREO	1
1.1 Conceptos generales de evaluación y rendimiento de los sistemas operativos	3
1.1.1 Necesidad de evaluar el rendimiento de los sistemas operativos	4
1.1.2 Necesidad de medir la carga del sistema operativo	5
1.1.3 Aspectos a tomar en cuenta para evaluar un sistema operativo.....	6
1.1.3.1 Percepción del rendimiento del sistema operativo por parte del usuario final.....	7
1.1.3.2 Percepción del rendimiento del sistema operativo por parte del administrador	8
1.1.3.3 Otras magnitudes relativas al comportamiento y rendimiento del sistema operativo.....	9

1.1.4	Parámetros que controlan el comportamiento del sistema operativo	10
1.2	Técnicas más comunes al evaluar un sistema operativo.....	12
1.2.1	Modelos analíticos.....	12
1.2.2	Modelos híbridos	13
1.2.3	<i>Benchmark</i>	13
1.2.4	Monitores.....	14
1.3	Definición de monitores como técnica de evaluación	14
1.3.1	Herramienta de monitoreo.....	15
1.3.2	Clasificación de los monitores.....	16
1.3.2.1	Clasificación de los monitores según el modo de su implementación	17
1.3.2.1.1	Monitores <i>software</i>	17
1.3.2.1.2	Monitores hardware	17
1.3.2.1.3	Monitores híbridos	18
1.3.2.2	Clasificación de los monitores por el mecanismo de activación	19
1.3.2.2.1	Monitores de eventos o acontecimientos	19
1.3.2.2.2	Monitores de muestreo	19
1.3.2.3	Clasificación de los monitores por la forma de manipular y presentar resultados	20
1.3.2.3.1	Monitores en tiempo real	20
1.3.2.3.2	Monitores <i>batch</i>	20
1.3.3	Combinación de las clasificaciones de los monitores.....	21
1.3.3.1	Estructura del monitor	22
1.3.3.2	Descripción de las funciones de los componentes de la estructura de un monitor	23

2 MONITORES PARA LA EVALUACIÓN DE LOS SISTEMAS OPERATIVOS	27
2.1 Terminología utilizada para el monitoreo del sistema operativo.....	27
2.1.1 Objetos	28
2.1.2 Contadores	28
2.1.3 Instancias.....	29
2.1.4 Umbrales	29
2.2 Identificación de los principales recursos del sistema operativo a monitorear	30
2.2.1 Memoria.....	31
2.2.2 Procesador	32
2.2.3 Discos	33
2.2.4 Redes	35
2.2.5 Otras áreas importantes que afectan el rendimiento	36
2.3 Objetivo de los componentes de los monitores para la evaluación de los sistemas operativos.....	37
2.3.1 Sistema de búsqueda y recopilación de datos.....	38
2.3.2 Analizador de datos	41
2.3.3 Intérprete de datos.....	41
2.4 Herramientas gráficas para el monitoreo de los sistemas operativos.....	42
2.4.1 Opciones de autoconfiguración	43
2.4.2 Alertas	44
2.4.3 El visor de sucesos	44
2.5 Contadores.....	45
2.5.1 Velocidad de operación	45
2.5.2 Detalle de recursos	46
2.5.3 Factor de utilización y disponibilidad	47
2.5.4 Contador de procesos	48

2.5.5	Control de flujo de la información	48
-------	--	----

3 AFINAMIENTO Y RENDIMIENTO, SOLUCIÓN A PROBLEMAS

DEL SISTEMA OPERATIVO	49
------------------------------------	-----------

3.1	Tareas del administrador de sistema	50
-----	---	----

3.2	Objetivos de una evaluación	52
-----	-----------------------------------	----

3.3	Análisis del afinamiento del sistema operativo	53
-----	--	----

3.3.1	El arte del afinamiento del sistema operativo	54
-------	---	----

3.3.2	Afinamiento y rendimiento del sistema	54
-------	---	----

3.3.2.1	Objetivos al momento de realizar un afinamiento	56
---------	--	----

3.3.2.2	Identificación de los recursos críticos	57
---------	---	----

3.3.2.3	Minimizando los requerimientos de los recursos críticos	60
---------	--	----

3.3.2.4	Modificación de la asignación de los recursos para reflejar la prioridad	62
---------	---	----

3.3.2.5	Repetición de los pasos del afinamiento del rendimiento	62
---------	--	----

3.4	Evaluación del rendimiento y afinamiento de la memoria	63
-----	--	----

3.4.1	Problemas de la memoria	63
-------	-------------------------------	----

3.4.1.1	Pérdida de la memoria	64
---------	-----------------------------	----

3.4.1.2	Paginación	64
---------	------------------	----

3.4.1.3	<i>Swapping</i> (alternar o intercambiar)	65
---------	---	----

3.4.2	Parámetros para medir y evaluar el uso de la memoria	65
-------	---	----

3.4.2.1	Páginas por segundo	66
---------	---------------------------	----

3.4.2.2	Bytes comprometidos	67
---------	---------------------------	----

3.4.2.3	Memoria virtual activa	67
---------	------------------------------	----

3.4.2.4	Porcentaje de uso del archivo de paginación	68
3.4.2.5	Bytes memoria no paginable	69
3.4.2.6	Memoria física disponible	69
3.4.3	Soluciones a los problemas del uso de la memoria.....	70
3.4.3.1	Administrador de la memoria.....	70
3.4.3.2	Tamaño de paginación y <i>swap</i>	71
3.4.3.3	Programas con uso de memoria intensivo.....	71
3.4.3.4	Páginas de memoria de archivos y datos	72
3.5	Evaluación del rendimiento y afinamiento del procesador o CPU	76
3.5.1	Problemas con el procesador	77
3.5.1.1	Programas mal diseñados	77
3.5.1.2	Cambios de contextos	78
3.5.2	Parámetros para evaluar y medir el uso del procesador.....	79
3.5.2.1	Interrupciones por segundo	79
3.5.2.2	Porcentaje de tiempo de interrupción	80
3.5.2.3	Tiempo de uso del procesador	80
3.5.2.4	Colas de procesos	81
3.5.2.5	Medición del uso del procesador por la ejecución de las aplicaciones	83
3.5.2.6	Identificación del uso intensivo del procesador por las aplicaciones	84
3.5.3	Soluciones a los problemas del procesador	85
3.5.3.1	Prioridad de ejecución	86
3.5.3.2	Tiempo de asignación del procesador	87
3.5.3.3	Ajuste de la carga del sistema.....	87
3.5.3.4	Memoria caché	88

3.6	Evaluación del rendimiento y afinamiento del disco	92
3.6.1	Problemas con los dispositivos de almacenamiento	92
3.6.1.1	Fragmentación	93
3.6.1.2	Lectura de archivos	93
3.6.1.3	Escritura de archivos.....	94
3.6.1.4	Contención	95
3.6.1.5	Atributos del sistema de archivo que afectan el rendimiento	96
3.6.1.5.1	Definición de los bloques de datos en el disco.....	97
3.6.1.5.2	Compresión de datos.....	97
3.6.2	Parámetros del disco a evaluar.....	98
3.6.2.1	Tiempo de utilización del disco.....	99
3.6.2.2	Tiempo de espera por disco.....	100
3.6.2.3	Transferencia de datos por segundo.....	100
3.6.2.4	Media en segundos por transferencia	101
3.6.2.5	Cola de actividades del disco.....	101
3.6.2.6	Tiempo por espera de lectura/escritura del disco.....	102
3.6.2.7	Desempeño de utilización del disco	103
3.6.2.8	Fragmentación del disco	104
3.6.2.9	Eficiencia de espacio y secuencialidad en el disco	104
3.6.3	Soluciones para el afinamiento de los discos.....	105
3.6.3.1	Embotellamiento por lectura o escritura al disco.....	106
3.6.3.2	Archivo de paginación	108
3.6.3.3	Utilización de adaptadores del disco.....	108
3.6.3.4	<i>Striping</i>	109

3.6.3.5	Sugerencias para el ajuste y progresión de la actividad del disco.....	109
3.6.3.6	Agregar más discos para mejorar el rendimiento.....	110
3.7	Evaluación del rendimiento y afinamiento de la red	116
3.7.1	Problema de la red	116
3.7.2	Parámetros para medir y evaluar el rendimiento de las redes	117
3.7.2.1	Paquetes recibidos	118
3.7.2.2	Colas de recepción y envío de paquetes.....	119
3.7.2.3	Colisiones	119
3.7.2.4	Fragmentación de paquetes	119
3.7.2.5	Redes <i>Ethernet</i>	120
3.7.2.6	Redes con protocolo IP (<i>Internet Protocol</i>).....	120
3.7.2.7	Protocolo UDP (<i>User Datagram Protocol</i>)	122
3.7.2.8	Protocolo TCP	123
3.7.3	Soluciones a los problemas de la red	124
3.7.3.1	Tamaño de los bloques de datos.....	126
3.7.3.2	Maximizando el rendimiento	127
3.7.3.3	Afinamiento de los adaptadores de colas de transmisión y de recepción	129
3.7.3.3.1	Colas de transmisión.....	130
3.7.3.3.2	Colas de recepción.....	131

4 CASO DE ESTUDIO DE EVALUACIÓN DEL RENDIMIENTO DE LOS SISTEMAS OPERATIVOS.....137

4.1	Análisis de la información de los comandos de línea, con monitores gráficos.....	138
4.2	Monitores gráficos para el análisis del sistema operativo	144

4.3	Análisis de la memoria.....	146
4.3.1	Consecuencias de poca disponibilidad de memoria física.....	147
4.3.2	Aspectos a monitorear para determinar la poca disponibilidad y problemas de la memoria.....	147
4.3.2.1	Memoria adicional	148
4.3.2.2	Memoria caché.....	150
4.3.2.3	Paginación de la memoria.....	153
4.3.2.4	Transferencia de páginas de archivo	155
4.4	Análisis del procesador o CPU	156
4.4.1	Consecuencias sobre problemas del procesador o CPU.....	157
4.4.1.1	Bajo rendimiento del procesador.....	158
4.4.1.2	Demanda excesiva de uso del procesador	158
4.4.2	Análisis de procesos y <i>threads</i> existentes.....	158
4.4.2.1	Cola de procesos	159
4.4.2.2	Distribución de actividades por el procesador	160
4.4.2.3	Utilización del procesador	162
4.4.2.4	Carga de los procesos	166
4.5	Análisis del disco	167
4.5.1	Consecuencias de un mal rendimiento del disco	168
4.5.2	Aspectos a monitorear por el bajo rendimiento del disco.....	169
4.5.2.1	Factor de utilización de los discos.....	169
4.5.2.2	Transferencia del disco	171
4.5.2.3	Lecturas y escrituras del disco	172
4.5.2.4	Tiempo de espera por disco.....	173
4.5.2.5	Fragmentación del disco	174

4.5.2.6	Resumen del rendimiento del disco.....	176
4.5.3	Embotellamiento por utilización de los discos.....	176
4.5.3.1	Longitud de espera por uso de los discos	177
4.5.3.2	Disponibilidad de almacenamiento en el disco	178
4.5.3.3	Excesivo <i>swap</i> del disco.....	180
4.6	Análisis de la red.....	181
4.6.1	Consecuencias de un mal rendimiento de la red	182
4.6.2	Aspectos a monitorear por el bajo rendimiento de la red	182
CONCLUSIONES		189
RECOMENDACIONES		191
BIBLIOGRAFÍA		193

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Estructura del monitor	22
2.	Estructura del sistema operativo	39
3.	Indicador de velocidad de operación	45
4.	Contador de recursos utilizados y disponibles	46
5.	Contador de recursos utilizados	46
6.	Contador de medición continua	47
7.	Contador de medición continua de forma volumétrica.....	47
8.	Contador de procesos	48
9.	Control de flujo de la información	48
10.	Áreas de detección y diagnóstico de los monitores.....	50
11.	Solicitud de servicio NFS.....	59
12.	Comportamiento de la memoria caché.....	141
13.	Rendimiento y disponibilidad de la memoria física y la memoria <i>swap</i>	142
14.	Monitores gráficos en ambiente Windows®	144
15.	Herramienta de monitoreo para sistema operativo <i>Spotlight® on Windows</i>	145
16.	Medición de rendimiento y disponibilidad de la memoria	146
17.	Utilización de recursos de los procesos activos	148
18.	Memoria virtual	149
19.	Eficiencia en la utilización de la memoria caché.....	150

20.	Deficiencia de rendimiento de la memoria caché.....	151
21.	Indicador del tamaño de memoria caché utilizada	152
22.	Páginas de memoria solicitadas que no se encuentran disponibles en el área de trabajo	154
23.	Transferencia de páginas de archivos del disco a la memoria.....	155
24.	Uso del procesador	156
25.	Cola de procesos y cantidad de procesos y <i>threads</i> en el sistema.....	159
26.	Porcentaje de atención de diferentes actividades del procesador	161
27.	Utilización del procesador	162
28.	Actividad de procesamiento atendido por el procesador.....	163
29.	Estadísticas de procesos en ejecución	164
30.	Seguimiento de un proceso determinado.....	165
31.	Unidades lógicas del disco en un sistema operativo.....	167
32.	Factor de utilización del disco	170
33.	Resumen de actividad de los discos.	170
34.	Tiempo de transferencia de los discos.....	171
35.	Escritura y lectura de discos en un rango de tiempo determinado.....	172
36.	Actividad de I/O reportado por segundo de disco.	173
37.	Espera de procesador por el disco.....	174
38.	Medición de la fragmentación, utilizando el <i>disk defragmenter</i> de Windows®.....	175
39.	Longitud de peticiones de disco en cola de espera.....	177
40.	Espacio ocupado y disponible en el sistema de disco, representado en gigabytes.....	179
41.	Actividad de lectura y escritura de disco por paginación y <i>swap</i>	180
42.	Estatus de descripciones de conexiones por red	182
43.	Monitoreo de un cuello de botella en la red	183

44.	Cuello de botella del adaptador 100Base-TX, tomado con la herramienta <i>Performance Monitor</i> de Windows®.....	184
45.	Descripción de opciones que detalla la herramienta de monitoreo <i>Distinct Network Monitor</i>	185
46.	Presentación de información registrada con la herramienta <i>Distinct Network Monitor</i> para el monitoreo de la red.....	187

TABLAS

I.	Funciones de los monitores <i>software</i> , <i>hardware</i> e híbrido	21
II.	Tareas de un administrador de sistema	51
III.	Clasificación de los recursos físicos y lógicos	58
IV.	Parámetros para la evaluación y rendimiento de la memoria.....	73
V.	Parámetros para la evaluación y rendimiento del procesador.....	89
VI.	Parámetros para la evaluación y rendimiento del disco	111
VII.	Tamaños de bloques para dispositivos y adaptadores de la red.....	126
VIII.	Parámetros para la evaluación y rendimiento de la red.....	132
IX.	Detalle de porcentajes disponibles en discos.....	179

LISTA DE SÍMBOLOS

Símbolo	Significado
\$	Indicador del símbolo del interprete de comandos del sistema operativo
K	Prefijo kilo
M	Prefijo mega
G	Prefijo giga

GLOSARIO

Arquitectura abierta	Arquitectura en la cual las especificaciones de diseño son publicadas y pueden añadirse más opciones.
Arquitectura cerrada	Arquitectura en la cual las especificaciones de diseño no son públicas y solamente pueden ser mejoradas por medio del fabricante.
<i>Background</i>	Término utilizado en informática para hacer mención a procedimientos y programas en ejecución en segundo plano.
<i>Backup</i>	Copia de seguridad
<i>Batch</i>	Proceso que ejecuta una secuencia de instrucciones de manera predefinida.
Bit	Acrónimo de <i>binary digit</i> (dígito binario), que adquiere el valor de 0 ó 1 en el sistema numérico binario.

Bitácora	Mecanismo que registra las actividades realizadas en un determinado momento por un proceso, dejando constancia de valores previos y posteriores.
Buffer	Memoria intermedia de manera temporal, donde se almacena datos hasta tener la oportunidad de completar su transferencia.
Byte	Unidad de información que consta de 8 bits.
Carga del sistema	La medida de cómo un <i>software</i> determinado está utilizando el <i>hardware</i> con una determinada combinación de programas.
Computadora	<p>Término utilizado para hacer referencia a uno o más equipos de cómputo a medir y evaluar las áreas de interés como memoria, disco, utilización del procesador, etc.</p> <p>El término computador es utilizado en términos generales, y puede ser un servidor de aplicaciones de propósito general o específico, o también llamado ordenador.</p>

Datagramas	Paquete de información que se transporta en la red con las direcciones de origen y destino.
Disponibilidad del sistema	Es la función del tiempo definida como la probabilidad de que el sistema esté trabajando correctamente y que esté disponible para realizar sus funciones en el instante considerado t . Incluye la posibilidad de que el sistema pueda haber estado averiado y posteriormente reparado.
Ethernet	Protocolo y estructura de red, para comunicación entre diferentes computadoras.
Fiabilidad del sistema	Es la función del tiempo definida como la probabilidad condicional de que el sistema trabajará correctamente a lo largo del intervalo de tiempo $[t_0, t]$, si trabajaba correctamente en el instante t_0 . En otras palabras es la probabilidad que el sistema trabaje correctamente a lo largo de un intervalo de tiempo dado. Se mide por la probabilidad de fallos por unidad de tiempo, o por el tiempo medio entre fallos
File system	Sistema de archivos, método para almacenar y guardar de manera ordenado información.

Gateway Puerta de comunicaciones de red, es una combinación de programa y de *hardware* que comunica dos tipos diferentes de red.

Hardware Parte física del computador.

Herramienta de monitoreo Herramienta utilizada para medir y evaluar parámetros en determinada área de aplicación, puede estar implementado por *software*, *hardware* o una combinación de ambos – monitor híbrido-

Existen monitores de sistemas operativos y monitores de aplicaciones.

Software Programa desarrollado para ejecutar un número determinado de instrucciones con cierta secuencia lógica, parte abstracta de una computadora.

Kernel Módulo central del sistema operativo, es la parte que se carga inicialmente a la memoria y es encargada de manejar la memoria, disco y funciones básicas iniciales de la computadora.

LAN	Red de área local, que consiste en un grupo de dos o más computadoras conectadas entre sí, generalmente por cable.
Mantenibilidad	Es la medida de la facilidad con que un sistema puede ser reparado después de un fallo. En términos cuantitativos es la probabilidad de que un sistema averiado pueda ser reparado y devuelto al estado operacional dentro de un período de tiempo determinado.
Memoria caché	<p>Memoria intermedia entre la memoria principal y el procesador que por sus características es muy rápida para leer y escribir.</p> <p>La memoria caché de disco, es parte de la memoria principal, que realiza una función similar pero a nivel del disco y la memoria principal.</p>
Overhead	Sobrecarga de información o tareas adicionales.
Performabilidad del sistema	Es una función del tiempo definida como la probabilidad de que el rendimiento del sistema estará por encima de un cierto nivel en un instante determinado.

Protocolo	Estándar que maneja la transmisión de datos en la red.
Quantum	Terminología utilizada en sistemas operativos para hacer referencia al tiempo que a un proceso le asignan el uso del procesador.
RAID	Acrónimo de <i>Redundant Array Independent Disk</i> (arreglo redundante de discos independientes), estructura de discos que trabajan de manera conjunta para ofrecer un mejor desempeño.
RAM	Acrónimo de <i>Random Access Memory</i> , (memoria de acceso aleatorio), siglas utilizadas para hacer referencia a la memoria principal de la computadora.
Router	Dispositivo que conecta dos dispositivos de área local.
Seguridad del sistema	Es la probabilidad de que el sistema esté realizando correctamente sus funciones o parado de forma tal que no perturbe el funcionamiento de otros sistemas ni comprometa la seguridad de las personas relacionadas con él.

Socket	Estructura de <i>software</i> que funciona como punto final de las comunicaciones dentro de un dispositivo de red.
Swap	Proceso interno de los sistemas operativos para disponer de mas memoria, almacenando información al disco.
Thread	Estado actual de ejecución de una instancia simple de un programa. Es un flujo simple secuencial de control de un proceso.
Throughput	Rendimiento de trabajo de un sistema por unidad de tiempo.
WEB	<i>www (world wide web)</i> , mecanismo proveedor de información electrónica para usuarios conectados en Internet.

RESUMEN

Considerando que los sistemas operativos son los administradores de los recursos de un computador y las empresas necesitan de las computadoras para la realización de muchas de sus actividades, estos deben de tener el mejor del rendimiento y desempeño posible. Por lo que es necesario monitorear y afinar constantemente los sistemas operativos para que sean eficientes en la administración de los recursos asignados y cumplir con las actividades programadas.

Cada sistema operativo cuenta con utilerías y comandos para evaluar el desempeño y rendimiento de cada uno de los recursos que administra. Pero debido a lo importante del afinamiento del sistema operativo, existen en el mercado diversas compañías de *software* que han desarrollado herramientas gráficas para la evaluación y afinamiento de sistemas operativos que además proporcionan ayuda utilizando una base de conocimiento y las mediciones registradas del rendimiento. Las herramientas de monitoreo proporcionan de manera oportuna y certera, consejos necesarios para modificar o corregir los parámetros de configuración del sistema operativo que permiten tener un mejor desempeño general.

Con este estudio se darán a conocer algunos de los aspectos más importantes al momento de realizar un afinamiento de los sistemas operativos, como la identificación de los parámetros de mayor importancia a evaluar.

OBJETIVOS

- **General**

Lograr el máximo rendimiento de los sistemas operativos, utilizando las herramientas nativas y herramientas disponibles en el mercado para la evaluación y rendimiento de sistemas operativos.

- **Específicos**

1. Adquirir los conocimientos básicos necesarios para llevar a cabo la medición del rendimiento de los sistemas operativos, estableciendo las variables que definen y caracterizan al sistema operativo y permiten controlar su comportamiento.
2. Comprender los términos utilizados en la configuración y monitoreo de los sistemas operativos. Y describir los conceptos básicos sobre monitores y herramientas de monitoreo, así como la estructura e información proporcionada para el análisis de afinamiento del sistema operativo.

3. Clasificar las diferentes áreas de los sistemas operativos que pueden ser puntos clave en el rendimiento y disponibilidad, para identificar y utilizar correctamente los distintos índices y valores proporcionados por las herramientas de monitoreo, en la medición de los sistemas operativos.

4. Utilizar los diferentes tipos de herramientas de monitoreo disponibles para sistemas operativos, identificando las principales cualidades y características de la herramienta de monitoreo así como su utilidad en las diferentes áreas de afinamiento de los sistemas operativos.

INTRODUCCIÓN

Los sistemas operativos trabajan con recursos y procesos asignados para realizar actividades solicitadas por diferentes tareas, y generalmente se requiere que estas actividades se realicen de manera rápida. En algunos casos el tiempo de respuesta depende mucho de la manera en que está configurado el sistema operativo, por lo que es necesario tener la mejor configuración según los recursos y la orientación específica de uso.

En el presente documento se determinan cuáles procesos y recursos pueden estar causando que el sistema en general responda inadecuadamente y las causas por las cuales se generan estos inconvenientes de rendimiento del sistema, se identifican qué parámetros se deben de medir, comparar, calcular y observar.

Para la realización de medición y evaluación de los parámetros se hace uso de las diferentes herramientas para monitoreo y comandos propios del sistema operativo.

Por último se expone una serie de mediciones de las áreas principales de afinamiento y la manera en que las herramientas de monitoreo proporcionan la información de manera fácil y entendible.

1 FUNDAMENTOS TEÓRICOS DE MEDICIÓN, EVALUACIÓN, Y HERRAMIENTAS DE MONITOREO

Existen diversos motivos por los cuales se desea evaluar el desempeño y rendimiento del sistema operativo, pero las principales razones son las que a continuación se describen

- Cuando se requiere de una nueva máquina para una tarea nueva y es necesario determinar qué sistema operativo es el más adecuado implementar.
- Cuando existe un nuevo sistema operativo que presenta mejores expectativas.
- Cuando se desea planificar la capacidad del sistema operativo actual, para tareas determinadas.
- Por último y la razón más frecuente, que es el de afinar o ajustar el sistema operativo por crecimiento de las aplicaciones y carga de la información.

Este último inciso, es el tema principal de estudio de este trabajo de graduación, pues comúnmente en toda empresa o institución la necesidad es maximizar la utilización de los recursos existentes, y en el ámbito informático, el afinamiento de los servidores es crucial para responder a las necesidades de procesamiento de información.

La necesidad de medir el rendimiento de un computador en términos del afinamiento del sistema operativo es determinar que esté funcionando de manera óptima para poder aprovechar al máximo los recursos disponibles de la computadora.

Las mediciones se deben de determinar por medio de datos sobre el rendimiento y comportamiento de las aplicaciones, en otras palabras, poder determinar cuáles son los cambios necesarios, para poder tener un mejor rendimiento, con los mismos recursos o bien, que recursos son indispensables ampliar para mejorar el rendimiento general del sistema operativo.

Para poder determinar los valores de los parámetros que afectan de manera drástica el rendimiento del sistema operativo se puede utilizar muchos métodos como uso de comandos propios del sistema operativo, utilización de *software* específico para medir y evaluar parámetros o bien, equipo de tipo *hardware* especial para medir los parámetros.

Cada método en la medición de los valores tiene sus ventajas y desventajas, el desarrollo de este trabajo está específicamente orientado a la ayuda de herramientas de ambiente gráfico de monitoreo, por su interpretación e información proporcionada en la evaluación y afinamiento de sistemas operativos.

1.1 Conceptos generales de evaluación y rendimiento de los sistemas operativos

La evaluación del rendimiento de los sistemas operativos o evaluación del desempeño, no es simplemente ampliar o agregar más recursos al computador, si bien este es un proceso muy empírico, se debe de destacar que no necesariamente el aumento de más recursos implica un aumento en el rendimiento del computador o una mejora sustancial y significativa, por otra lado, este proceso no implica ningún mecanismo de evaluación, como tampoco proporciona un índice de mejora.

Por el contrario, un método de evaluación apropiado proporciona información sobre qué áreas de la computadora necesitan ser revisadas, para determinar el rendimiento, con base en estadísticas y niveles de valores aceptables para determinados parámetros del sistema operativo.

También es importante mencionar que se debe de registrar el rendimiento antes y después de los cambios sobre el sistema operativo de manera tal que pueda existir un análisis sobre la mejora en el rendimiento.

Kant (1980,24), describe que un sistema puede ser evaluado en términos de que tan bien cumple con las objetivos y necesidades, esperados por los usuarios en referencia a su desempeño; y para poder medir esto, dice que se deben de tomar en cuenta como responde el sistema a las necesidades de las aplicaciones, si satisface eficientemente los requerimientos.

Este mismo autor describe también que toda medida de rendimiento debe de tratar con tres objetivo básicos que son:

- Qué tan rápida, una tarea puede ser cumplida.
- Qué tan bien puede trabajar el sistema con fallas y situaciones anormales.
- Qué tan eficientemente utiliza los recursos el sistema operativo.

1.1.1 Necesidad de evaluar el rendimiento de los sistemas operativos

Para poder mejorar la rapidez con que las tareas del sistema operativo son realizadas, es necesario determinar como esta inicialmente el rendimiento y con qué recursos cuenta, como también cómo los utiliza, para hacer esto es necesario poder evaluar el rendimiento y que mejor que una herramienta como ayuda.

Si bien en el mercado informático existen varios tipos de sistemas operativos de arquitectura cerrada o arquitectura abiertas, siempre cuentan con áreas y funciones similares, como controlar los procesos, coordinar y manipular el *hardware* del computador, organizar los archivos en diversos medios de almacenamiento y gestión de errores de *hardware* y pérdida de datos. Estas funciones son las que deben de ser monitoreadas principalmente, para medir y evaluar el rendimiento de los sistemas operativos.

La necesidad de medir un sistema operativo, con respecto al rendimiento y capacidad, es para determinar de manera cuantitativa y cualitativa cómo un *software* o un conjunto de programas de computación e inclusive el propio *software* del sistema está utilizando el *hardware* y la disponibilidad de recursos del sistema de cómputo para llevar a cabo las tareas que se le han encomendado.

1.1.2 Necesidad de medir la carga del sistema operativo

El comportamiento del sistema operativo es muy dependiente de la carga que procesa, pues en muchos casos, ésta crece de manera no programada o planificada, ya sean nuevas aplicaciones que inicialmente no estaban previstas o bien un aumento sustancial de los datos iniciales a manipular, que por otro lado afectan el rendimiento.

La necesidad de medir la carga del sistema operativo es para determinar el grado en que se están aprovechando los recursos del sistema, como por ejemplo, la memoria, uso del procesador, uso del disco, etc., y determinar el funcionamiento general, disponibilidad y funcionalidad, el cual al poder realizar las mediciones de carga del sistema, se determinará si el funcionamiento es el correcto, es decir, el esperado.

Otro de los beneficios es determinar qué factores impiden un funcionamiento adecuado, como pueden ser los cuellos de botella, frecuencia de solicitud y uso de un recurso determinado no disponible, disponibilidad y tiempos de descanso (tiempo no utilizado) de determinados recursos, etc. Kant (1980,24) dice, que al medir un sistema, también se puede determinar que tan bien trabaja en fallas determinadas.

Por consiguiente, para poder obtener un mejor desempeño se deben identificar los componentes principales que afectan el rendimiento del sistema operativo como: memoria física, medios de almacenamiento secundario (disco), procesador, acceso a dispositivos de red; y seguidamente identificar los parámetros y valores que no están de acuerdo con los niveles aceptables.

Para poder determinar que el valor de un parámetro determinado está en los valores adecuados o fuera de los rangos aceptables, es necesario contar con una herramienta o método el cual permita medir e interpretar adecuadamente las mediciones de la carga del sistema y proporcione información adicional la cual permita comprender o identificar que niveles de valores son los adecuados para cada uno de los elementos y parámetros que se está midiendo y evaluado, según las características y uso del sistema operativo. Por lo cual es indispensable contar con una base de conocimiento de estas mediciones y que sea empleada por una herramienta de *software* de fácil uso para los administradores de sistemas operativos.

1.1.3 Aspectos a tomar en cuenta para evaluar un sistema operativo

Para determinar cuando es el tiempo indicado para evaluar el comportamiento de un sistema operativo hay que hacer referencia a unas medidas cuantitativas y parámetros que van a caracterizar el comportamiento del *hardware* y *software* del computador que hacen referencia a cómo el usuario final hace uso del sistema, ya sea por el uso de comandos o por el uso de aplicaciones que no responden adecuadamente.

Otro indicador es el administrador del sistema, quien administra y distribuye la carga, y es por medio de estos dos usuarios principalmente (usuario de visión externa y usuario de visión interna), que se determina cuando es necesario realizar una revisión (medición o evaluación) en determinado momento, tomando como base la percepción que tienen del rendimiento.

Los parámetros y magnitudes en que se basa la percepción de los usuarios están relacionados con tres tipos de medidas correspondientes a.

- Tiempo de respuesta
- Utilización de recursos o dispositivos
- Trabajos realizados por el sistema o componentes del mismo

1.1.3.1 Percepción del rendimiento del sistema operativo por parte del usuario final

Estos parámetros o variables del rendimiento de operación, son los que él usuario percibe, con base a la interacción y rendimiento del sistema¹.

- Productividad o *throughput*: es la cantidad de trabajo ejecutado por unidad de tiempo.
- Capacidad: la máxima cantidad de trabajo útil que se puede realizar por unidad de tiempo
- Tiempo de respuesta: el tiempo transcurrido entre la entrega de un trabajo o una transacción al sistema y la recepción del resultado o la respuesta.

¹ Estos parámetros son conocidos también como criterios para realizar el afinamiento de sistemas operativos.

1.1.3.2 Percepción del rendimiento del sistema operativo por parte del administrador

Estos parámetros son las variables internas del sistema, las cuales dependen en parte del diseño y tiempo de respuesta de mecanismos propios del *hardware*, frecuencia de utilización, como también la interacción del sistema operativo con dicho *hardware*. Pueden ser mejoradas con forme el avance de tecnología y adquisición de mayor cantidad de recursos de *hardware*, y son administradas por la persona encargada del sistema (administrador del sistema), quien tiene a su cargo medir y ajustar dichos parámetros.

Los principales parámetros internos del sistema son

- Factor de utilización de un componente, que es el porcentaje de tiempo durante el cual un componente del sistema informático (CPU, canal, dispositivo de E/S, etc.) ha sido realmente utilizado.
- Solapamiento de componentes, es el porcentaje de tiempo durante el cual dos o más componentes del sistema han sido utilizados simultáneamente.
- *Overhead*, es el porcentaje de tiempo que los distintos dispositivos del sistema (CPU, discos, memoria, etc.) han sido utilizados en tareas del sistema no directamente imputables a alguno de los trabajos en curso.
- Factor de carga de multiprogramación, es la relación entre el tiempo de respuesta de un trabajo en un determinado entorno de multiprogramación y su tiempo de respuesta en monoprogramación.

- Factor de ganancia de multiprogramación, es la relación entre el tiempo total necesario para ejecutar un conjunto de programas secuencialmente de multiprogramación y en multiprogramación.
- Frecuencia de fallo de página, es el número de fallos de página que se producen por unidad de tiempo en un sistema de memoria virtual paginada.
- Frecuencia de *swapping*, que el número de programas expulsados de la memoria por unidad de tiempo a causa de falta de espacio o con el fin de permitir su reorganización para recuperar espacio en ella o para disminuir la paginación.

1.1.3.3 Otras magnitudes relativas al comportamiento y rendimiento del sistema operativo

Hay otras medidas relacionadas con el comportamiento del sistema pero no directamente con el rendimiento, que en muchos casos también es importante tenerlas en cuenta y que simplemente se mencionan a continuación.

- Disponibilidad del sistema: es la probabilidad que indica con que frecuencia el sistema está disponible para realizar las actividades requeridas y que trabaje de manera correcta.
- Fiabilidad del sistema: proporciona una probabilidad que en un determinado momento que se realice una tarea, el sistema operativo trabaje de manera correcta.

- Seguridad del correcto funcionamiento del sistema: es la probabilidad de que el sistema esté trabajando correctamente o bien que esté parado sin afectar el funcionamiento de otros sistemas.
- Alto rendimiento del sistema: es la probabilidad de que el sistema proporcione de manera continua un nivel aceptable de funcionamiento.
- Mantenibilidad: es la probabilidad con que un sistema operativo, después de un fallo, pueda ser reparado en un tiempo determinado.

1.1.4 Parámetros que controlan el comportamiento del sistema operativo

Hasta este punto, se sabe porque se debe medir y los aspectos a tomar en cuenta, ahora se tratarán de los parámetros que afectan el comportamiento del sistema operativo, es decir magnitudes para controlar el comportamiento, cuando el rendimiento no es el esperado o el deseado.

Las modificaciones que se pueden introducir para mejorar ese comportamiento pueden hacerse a niveles que influyen en el comportamiento del mismo mediante

- Ajuste de los parámetros
- Modificación de las políticas de gestión
- Equilibrio de la distribución de cargas.
- Modificación o sustitución de componentes *hardware* del sistema.
- Modificación de los programas.

El desarrollo de este trabajo de graduación se enfoca en el aspecto de identificación y ajuste de los parámetros del sistema operativo. La modificación de las políticas de gestión es la aplicación como resultado del estudio del equilibrio de la distribución de cargas de manera escueta por medio de técnicas de simulación y uso de herramientas *benchmark*, con el fin de comprender la necesidad de estudiar el comportamiento de la carga en el sistema operativo.

La modificación o sustitución de componentes de hardware del sistema operativo se enfoca a determinar qué componentes *hardware* están sobrecargados por lo que es necesario cambiarlos por uno con mejor tecnología o más eficiente o bien incrementar la cantidad de estos componentes.

La modificación de programas o aplicaciones es otro tema aparte, aunque también es un área muy amplia, que debería de ser tratada en otro trabajo de graduación, pues incluye un área grande de investigación y técnicas de programación para hacer más eficaz el *software*, como también de políticas de afinamiento y escrituración de *software*.

1.2 Técnicas más comunes al evaluar un sistema operativo

Se denominan técnicas de evaluación a los métodos y herramientas que permiten obtener los índices del rendimiento de un sistema que está ejecutando, con una carga de sistema dada y con unos valores determinados de parámetros del sistema.

Las técnicas que se mencionan en este trabajo de graduación son unas de la muchas técnicas utilizadas para medir y evaluar el rendimiento de los sistemas operativos, claro está, cada una tienen sus pro y contra que dependen de los recursos, disponibilidad y tipo de sistema a medir. Se mencionan como referencias, aunque el estudio se orientaría al uso de herramientas de monitoreo de sistemas operativos, aquí se describen brevemente para identificar algunas de las opciones disponibles.

- Modelos analíticos
- Modelos híbridos
- *Benchmark*
- Monitorización

1.2.1 Modelos analíticos

Los modelos analíticos están basados fundamentalmente en la teoría de colas, técnicas de simulación o técnicas analíticas. El modelo analítico es ideal cuando se trata de evaluar el comportamiento de un sistema en el que hay algún elemento (*hardware* o *software*) que no está instalado.

La limitación de los métodos analíticos es su incapacidad para tratar determinadas estructuras y comportamientos de las colas que existen en los sistemas informáticos. Los métodos de simulación no tienen estas limitaciones, pero, en general son mucho más caros que los analíticos tanto en tiempo de cálculo como en esfuerzo de puesta a punto.

No obstante, la principal dificultad de esta herramienta reside en la obtención de datos lo suficientemente precisos para ejecutar el modelo y obtener resultados con el grado de aproximación que se exige.

1.2.2 Modelos híbridos

Los modelos híbridos, consisten en subdividir modelos complejos en el que cada submodelo representa aspectos importantes del sistema a medir.

La principal característica de esta técnica es que cada submodelo puede ser tratado, utilizando métodos diferentes, como métodos analíticos, uso de simulación, etc.

1.2.3 Benchmark

Es un método bastante frecuente para comparar sistemas informáticos frente a una carga característica de una instalación concreta, efectuándose la comparación, básicamente, a partir del tiempo necesario para su ejecución. Generalizando se puede considerar como la medición del comportamiento sobre un prototipo. Variantes de este método se usan para evaluar la potencia relativa de un sistema a lo largo de su ciclo de vida, para contrastar monitores y para validar modelos.

Las principales dificultades que se plantean al utilizar un *benchmarking* es poder determinar la carga característica, de forma que sea suficientemente reducida para ser manejable y suficientemente extensa para ser representativa; y valorar el aprovechamiento que hacen los programas de las peculiaridades de los distintos sistemas.

1.2.4 Monitores

Los monitores son las herramientas de medición que permiten seguir el comportamiento de los principales elementos de un sistema informático cuando éste se halla sometido a una carga de trabajo determinada por los usuarios del sistema.

Pueden existir monitores de tipo *hardware*, *software* o mixtas. Aunque su objetivo es la medición de la carga, se les denomina monitores ya que, debido a la imposibilidad de reproducir situaciones con la carga real, estos instrumentos hacen un seguimiento de lo que sucede en el sistema, es decir, lo monitorean.

1.3 Definición de monitores como técnica de evaluación

Los monitores comúnmente se utilizan para medir, pero también conllevan funciones adicionales como poder describir dónde y cómo poder optimizar los recursos que se está monitoreando, describir la utilización de los recursos, y proporcionar criterios de afinamiento a través de ajustes de parámetros y la planificación del crecimiento y expansión necesaria para soportar futuras cargas.

Existen varias definiciones sobre los monitores, de las cuales se mencionan algunas y luego se hace la definición del término de herramienta de monitoreo utilizada en este documento.

Jain (1980,30) describe los monitores como “herramientas utilizadas para observar las actividades del sistema en general” y describe sus funciones como, “observar el rendimiento del sistema, recolectar estadísticas de rendimiento y analizar datos y desplegar los datos analizados”.

Puigjaner (sa, 29) describe los monitores como “herramienta utilizada para observar la actividad de un sistema informático mientras es utilizado por sus usuarios, recopilando información estadística de los objetos (componentes de interés del sistema a monitorear), analiza los datos recogidos y presenta los resultados de manera legible y entendible. En muchos casos es capaz de sugerir cambios en la configuración del sistema para llevar a cabo un mejor desempeño del proceso del sistema.”

1.3.1 Herramienta de monitoreo

Una herramienta de monitoreo, es un programa que permite observar la carga del sistema, de manera general como individual en cada uno de los componentes, de una manera legible y fácil de entender (ya sea con gráficas, dibujos, números y colores) para comprender los cambios necesarios a realizar para mejorar el rendimiento, identificando las variables y proporcionando sugerencias sobre las mejoras a realizar. Crea además, un historial del comportamiento para futuros análisis y repetición del desempeño, tomando en cuenta que la carga que aporta el propio monitor es mínima o despreciable en las medidas.

Esta definición describe específicamente la palabra programa, pues aunque existen diferentes tipos de clasificación e implementación de monitores, el estudio se enfoca específicamente en el uso de monitores *software*, pero como se describe más adelante, por medio de combinación de otros conceptos sobre implementación de los monitores se realizar una especie de mezcla y obtener lo mejor de las clasificaciones, para hacer una herramienta más útil y versátil para el propósito.

1.3.2 Clasificación de los monitores

Los monitores están divididos en tres clasificaciones que tienen que ver en la forma en que se implementan, métodos de recopilación de información y presentación de resultados.

Clasificación según el modo de implementación

- Monitores *software*
- Monitores *hardware*
- Monitores híbridos

Clasificación según su mecanismo de activación

- Monitores de eventos o acontecimientos
- Monitores de muestreo

Clasificación según su forma de manipular y presentar los resultados

- Monitores en tiempo real
- Monitores *batch*

1.3.2.1 Clasificación de los monitores según el modo de su implementación

La clasificación por medio de su implementación está basada en la manera en que se puede llevar a cabo la monitorización de los objetos a medir o cuantificar.

1.3.2.1.1 Monitores *software*

Son programas o ampliaciones del sistema operativo que acceden al estado del sistema, informando al usuario (ya sea de forma continua, o cuando éste lo requiere) sobre dicho estado. Los monitores *software* son los más fáciles de encontrar y son utilizados de forma general para estudiar el comportamiento de un sistema.

El código que necesitan los monitores de software para llevar a cabo su tarea, se puede implementar en tres formas distintas

- Adición de un nuevo programa
- Modificación del *software* que se va a medir
- Modificación del sistema operativo

1.3.2.1.2 Monitores *hardware*

Son dispositivos electrónicos que se conectan a determinados puntos del sistema, donde se encargan de detectar determinados niveles o señales eléctricas que caracterizan la evolución del sistema.

La principal característica de estos es que son externos al sistema que van a medir lo que implica que:

- No utilizan recursos del sistema que se va a monitorizar.
- No producen interferencias con éste.
- Son muy rápidos.

Por tanto, en un principio parece que sean mejores que los monitores *software*, pero tienen algunas desventajas frente a éstos que son

- Más difíciles de instalar.
- Existen magnitudes a las que sólo se puede acceder desde el *software*.
- Requieren, para su operación y para el análisis de sus resultados, de personal más especializado.
- Pueden interactuar a nivel eléctrico con el sistema que se va a monitorizar, provocando perturbaciones que resulten en un funcionamiento anómalo del sistema monitorizado.
- No se deben utilizar nunca en medidas sobre sistemas críticos en funcionamiento real.

1.3.2.1.3 Monitores híbridos

Son una combinación de las técnicas de monitores *software* y monitores *hardware*, intentando combinar las ventajas de uno y otro.

Una forma de implementar este tipo de monitores es que se debe de añadir sobre el sistema operativo una parte muy pequeña de código que envía información al monitor *hardware* que está observando de forma continua la ejecución del objeto a monitorear, y claro está, muchas veces no se cuenta con el código fuente o bien, se utilizan diferentes funciones y procedimientos de acceso al núcleo del sistema operativo desarrollado por *software* de manera específica por lo cual se requiere de diferentes mecanismos para su instalación.

1.3.2.2 Clasificación de los monitores por el mecanismo de activación

Esta clasificación de los monitores depende de la manera en que se activa o toma la decisión de recopilar información.

1.3.2.2.1 Monitores de eventos o acontecimientos

Son aquellos que se activan por la aparición de ciertos eventos. Si el evento se da con frecuencia, la sobrecarga que producen es elevada. Este evento puede ser también una decisión externa del responsable del sistema.

1.3.2.2.2 Monitores de muestreo

Los monitores de muestreo son aquellos que se activan a intervalos de tiempo fijo o aleatorio, mediante interrupciones de reloj. La frecuencia de muestreo viene determinada por la frecuencia de aparición del estado que se desea analizar y por la resolución deseada.

1.3.2.3 Clasificación de los monitores por la forma de manipular y presentar resultados

Esta clasificación está orientada según cómo trabaja el componente del monitor encargado de presentar la información analizada, puede variar su implementación, según la necesidad de rapidez y frecuencia de presentación de datos, que el usuario final (administrador del sistema operativo o analista) requiera por la orientación que tiene el sistema operativo.

1.3.2.3.1 Monitores en tiempo real

Los monitores de tiempo real constan de un módulo analizador que procesa los datos recogidos a medida que los recibe. En consecuencia, el módulo de proceso debe ejecutarse de forma continua, o con gran frecuencia, para poder presentar resultados parciales al analizador que informará de la evolución del sistema.

1.3.2.3.2 Monitores *batch*

Difieren de los monitores de tiempo real, pues éstos recogen toda la información y, una vez terminado el periodo de recolección, la misma es procesada.

1.3.3 Combinación de las clasificaciones de los monitores

Las tres clasificaciones sobre los monitores, pueden utilizarse de forma conjunta, así pues, un monitor puede clasificarse como *software*, conducido por muestreo y con análisis en *batch* de los resultados. La clasificación según el tipo de interfase es la más habitual.

En la siguiente tabla muestra cómo se realizan cada una de las funciones los monitores software, hardware e híbrido.

Tabla I. Funciones de los monitores *software*, *hardware* e híbrido

Función	Monitor <i>software</i>	Monitor <i>hardware</i>	Monitor híbrido
Instrumentación	Trampas	Sondas	Trampas
	Interceptación	Conexión permanente	Sondas
	Muestreo		
Selector	Programa	Programa	Programa
		Lógica	Lógica
Procesador	Programa	Lógica	Lógica
		Programa	Programa
Registrador	Memoria	Contador	Contador
	Disco	Memoria	Memoria
		Disco	Disco
Analizador	Programa	Programa	Programa

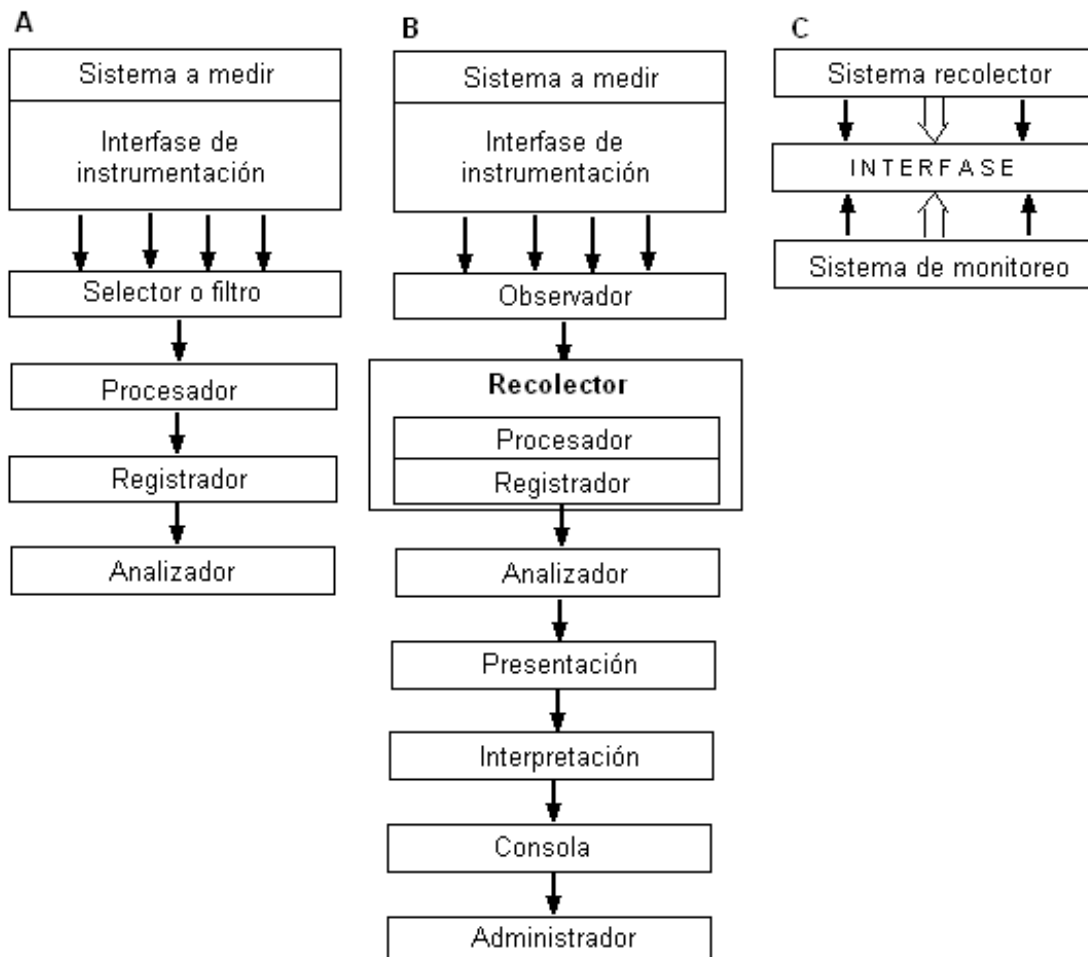
Fuente: Ramón Puigjaner y otros. **Evaluación y explotación de sistemas informáticos**, pág. 38.

Aquí solamente se detallan algunas de las muchas funciones que puede tener un monitor. Otras funciones que pueden tener dependen de la orientación y de las características propias del sistema operativo.

1.3.3.1 Estructura del monitor

La estructura del monitor es la parte conceptual (esquema conceptual) del monitor, en esta parte se detalla qué funciones o módulos contiene el monitor. Describe además la finalidad de cada uno de los componentes y el alcance que tendrá el monitor.

Figura 1. Estructura del monitor



Adaptado: Ramón Puigjaner y otros. **Evaluación y explotación de sistemas informáticos**, página. 36

La estructura del monitor que se ilustra en la página anterior, en el inciso A describe un monitor sencillo, tal como lo expresado por Puigjaner (sa, 42), que no incluye elementos como la consola, funciones de presentación o una unidad de administración sobre las funciones totales del monitor, pero es funcional y excluye estas funciones en la implementación del monitor.

La estructura del monitor B es de un monitor para un sistema distribuido, que contiene más elementos para configurar y manejar la información, esto es funciones de presentación, consola y administración, tanto del monitor como de análisis de la información recopilada.

La última estructura, del monitor C, es típicamente la estructura de un monitor híbrido (combinación de *software* y *hardware*), que detalla como un sistema de recolector a la parte de *hardware*, puede variar según el componente *hardware* que se desea monitorear, la interfase, que es en sí el punto de encuentro entre el monitor por *software*, que solicita y procesa los datos recolectados y proporcionados por el monitor *hardware*.

1.3.3.2 Descripción de las funciones de los componentes de la estructura de un monitor

Cada componente de los monitores tiene definida una función específica, algunas veces están bien identificadas y otras son un poco difíciles de identificar, a continuación se presenta un resumen de lo que hace y significa cada uno de los componentes en la estructura de los monitores.

- Sistema a medir: es el sistema a evaluar, en este caso es el sistema operativo del cual se realiza una evaluación del rendimiento, es el objetivo de utilización del monitor.
- Interfase de instrumentación: es la conexión entre el monitor y el sistema operativo, en este punto es donde se incluyen todos aquellos elementos que permiten ser monitoreados de lado del sistema operativo, tales como, por ejemplo, secciones de código incluido en el sistema operativo, temporizadores, puntos de sondeo, etc.
- Selector o filtro: es el elemento selector o de filtrado que permite una captura selectiva de las actividades de interés. Esto le permite al usuario orientar la medida hacia los aspectos específicos que desea explorar. Además, tiene por misión generar aquellas variables que no teniendo existencia física directa en el sistema, queden generarse en función de otras que son directamente observables.
- Procesador: es el elemento que realiza las comprobaciones de los elementos del sistema que van a ser analizados o medidos.
- Registrador: es todo dispositivo utilizado por el monitor para resguardar y registrar los resultados de las comprobaciones, funciona como una bitácora de los eventos realizados en el sistema en determinado momento.

- **Analizador:** es el elemento que realiza un análisis e interpretación detallado de las comprobaciones realizadas por el procesador y almacenadas en el registrador.
- **Presentación:** es la parte del monitor en el cual se presenta la información al usuario, ésta puede ser por datos tabulados, gráficas de líneas o pies, o bien una combinación de dibujos / caracteres especiales que representan una función específica, flujos de transferencia de datos para representar velocidades de transferencia como también combinación de colores para representar el grado de eficiencia y correcto funcionamiento de los componentes.
- **Interpretación:** corresponde en algunos tipos de monitores al significado de la forma de presentación de los datos, algunos incluyen en esta parte el afinamiento, el cual luego de procesar la información puede detallar una lista de sugerencias para el afinamiento y corrección de las áreas de problemas.

Algunas veces se pueden tener varias fases combinadas en la estructura del monitor, o bien dos o más fases pueden trabajar en paralelo, tal como por ejemplo, la fase de análisis e interpretación se realizan en paralelo con la detección y captura de eventos, mientras que otras veces se hace después de recoger toda la información.

2 MONITORES PARA LA EVALUACIÓN DE LOS SISTEMAS OPERATIVOS

En este capítulo, se define lo que hay que monitorear del sistema operativo y cómo se identifican a nivel de los monitores cada uno de los recursos de la computadora, para determinar que está funcionando correctamente y que necesita un ajuste para tener un mejor desempeño.

Se describirán brevemente los aspectos internos que necesitan ser monitoreados y principales características de estos con respecto al funcionamiento general del sistema operativo.

2.1 Terminología utilizada para el monitoreo del sistema operativo

Para poder identificar las áreas necesarias a ser monitoreadas en todo sistema operativo, primeramente se realiza una clasificación de los componentes que proporcionan información del rendimiento del sistema operativo y terminología utilizada para la identificación de los datos proporcionados por medio de los objetos, contadores e instancias.

2.1.1 Objetos

Un objeto de rendimiento (también llamado componente de rendimiento) es cualquier recurso, programa o servicio que se puede medir. Con esto es posible utilizar los monitores, registros y alertas para seleccionar que objetos se desea evaluar.

Un objeto de rendimiento genera información de su desempeño, cuando realiza alguna actividad en el equipo.

2.1.2 Contadores

Los contadores son componentes dentro de un objeto, estos representan información respecto al funcionamiento del objeto. Generalmente por cada objeto, existen muchos contadores asociados, de manera que se puede determinar el rendimiento con los valores asociados a los contadores. Cada vez que un objeto es utilizado, genera información susceptible de evaluación, es decir los contadores correspondientes se actualizan.

Los contadores asociados a un objeto pueden incrementarse por medio de programas, que generalmente son para medir el rendimiento sobre dicha área de un objeto que no cuenta con un parámetro o un valor específico de medición o bien como resultado de la operación de uno o más contadores.

2.1.3 Instancias

Se trata de una única ocurrencia de varios objetos de rendimiento del mismo tipo en el sistema operativo. Si un objeto en particular tiene múltiples instancias, se puede hacer un seguimiento de las estadísticas para cada instancia añadiendo un nuevo contador para cada una. También se puede añadir un contador para hacer un seguimiento de todas las instancias a la vez.

Por ejemplo, si se cuenta con dos procesadores (objeto procesador), se tienen dos instancias del objeto procesador. Si se tienen dos discos físicos, se tendrán dos instancias del objeto disco físico, en cambio existen algunos tipos de objetos que no tienen instancias, como por ejemplo, la memoria.

2.1.4 Umbrales

Los umbrales son en sí el rango de valores o el valor máximo en los cuales se puede clasificar que determinado objeto está en los límites aceptables de su rendimiento.

Por ejemplo, si la cola de procesos en espera es de dos unidades (dos procesos en la cola), está en el límite del umbral máximo óptimo de procesos en espera. Pero también puede ser los valores de los umbrales en términos de porcentaje, como en el caso de espacio del disco, si este tiene un espacio libre de 10 por ciento, está en su umbral mínimo permitido de espacio libre disponible.

El desarrollo de este trabajo de graduación pretende dejar establecidos algunos criterios para identificar qué valores debería de tener estos umbrales para los objetos principales a analizar y qué aspectos pueden ser corregidos para nivelar los valores de los contadores dentro de los umbrales aceptables.

2.2 Identificación de los principales recursos del sistema operativo a monitorear

Es importante identificar que componentes de *hardware* y *software* están instalados y qué reconoce el sistema operativo, las aplicaciones que están instaladas, quién las utiliza, cuál es la carga, e incluso la función del computador en la infraestructura de la red, para poder llevar a cabo un monitoreo, También es necesario establecer una referencia de rendimiento o desempeño que indique el modo en que el sistema utiliza los recursos durante los periodos de actividad normal de operación, pues sin este parámetro de medición, no es posible determinar el correcto o buen funcionamiento del sistema.

Los principales objetos que deben de ser analizados y tener referencias de rendimiento de trabajo y carga son:

- uso de memoria
- procesador
- discos (esto incluye archivos de paginación, uso y acceso del disco, caché del disco, en general dispositivos de almacenamiento secundario)
- interfase de red.

Estos cuatro objetos (aunque el sistema dispone de muchos más objetos, son los predeterminados para medir el desempeño del sistema) están estrechamente relacionados y cada uno de ellos puede estar afectado por el mal desempeño de otro de estos objetos, o bien puede causar que otro objeto funcione inadecuadamente o con un nivel bajo de rendimiento al esperado. Por lo que es importante realizar los cambios para la mejora del sistema, uno a la vez y verificar con base en los valores de referencia recopilados inicialmente, si existe una mejora con los cambios realizados.

Los principales aspectos de rendimiento que se deben de supervisar en estos objetos son

- Uso
- Cuellos de botella o problemas con los objetos
- Pérdidas de información
- Rendimiento

2.2.1 Memoria

La memoria del computador son semiconductores, que permiten almacenar y recuperar datos de manera rápida que se encuentra conectado directamente al procesador. En la memoria se colocan las aplicaciones de los usuarios y datos tanto para control interno de las tareas y programas, como datos de información. Por la velocidad de lectura y escritura es un recurso muy utilizado, pero el almacenamiento es de manera temporal y el espacio suele ser muy pequeño con respecto a almacenamientos secundarios, por lo que generalmente existen contenedores por la utilización de este recurso.

Cuando la memoria disponible a utilizar es escasa, habitualmente existen embotellamientos de memoria, los cuales se presentan afectando otros objetos como sobrecarga del procesador o bien un deficiente desempeño del disco, por lo que es necesario verificar el comportamiento de la memoria antes de estos objetos mencionados.

Existen varios índices de comportamiento de la memoria, como páginas por segundo, cantidad de bytes disponibles, porcentaje de memoria utilizada, etc., que son útiles para determinar el rendimiento de objeto memoria. Un mejor indicador del embotellamiento de memoria es una tasa alta y constante de errores de página de *hardware*, los cuales se producen cuando un programa no encuentra en la memoria física o memoria principal los datos que necesita y tiene que ir a recuperarlos a medios de almacenamiento secundario (memoria secundaria), como un disco el cual tiene un rendimiento inferior a la memoria.

2.2.2 Procesador

El procesador, también llamado microprocesador, es un circuito electrónico que actúa como unidad central de proceso de un ordenador, proporcionando el control de las operaciones de cálculo.

El microprocesador internamente contiene dispositivos de memoria, para almacenar registros de información de las instrucciones a ejecutar, registros de valores como registros de estados como resultado de las operaciones y del procesador.

El procesador no es capaz de albergar en su memoria la gran cantidad de información de instrucciones y datos de programa, por lo cual hace uso de la memoria para realizar sus funciones, muchas veces, la restricción de la memoria hace que el procesador presente problemas de rendimiento, es decir, que para determinar si el sistema operativo sufre de un embotellamiento de procesador, se debe determinar primero que el sistema operativo no sufre un embotellamiento de memoria.

El embotellamiento del procesador o del CPU, se produce únicamente cuando el procesador está tan atareado que no puede responder a las peticiones. Los síntomas de esta situación se encuentran en los elevados índices de actividad del procesador, grandes colas de forma frecuente y un mal tiempo de respuesta de las aplicaciones. Las causas habituales de estos síntomas, son por lo general las aplicaciones y controladores dependientes del procesador y las interrupciones constantes, provocadas por componentes mal diseñados de los subsistemas de la red o del disco.

2.2.3 Discos

Un dispositivo de almacenamiento, en ordenadores o computadoras, es todo aparato que se utilice para grabar los datos de la computadora de forma permanente o temporal. Una unidad de disco, es un dispositivo de almacenamiento. A veces se dice que una computadora tiene dispositivos de almacenamiento primarios (o principales) y secundarios (o auxiliares). Cuando se hace esta distinción, el dispositivo de almacenamiento primario es la memoria de acceso aleatorio (RAM) de la computadora, un dispositivo de almacenamiento permanente pero cuyo contenido es temporal.

El almacenamiento secundario incluye dispositivos de almacenamiento más permanentes, como unidades de disco y de cinta, las cuales tienen la característica de ser unidades de almacenamiento permanente de gran capacidad, pero con menor rapidez de acceso a la información.

Los dispositivos de almacenamiento secundario, pueden tener diferentes estructuras y arquitecturas en su implementación para almacenar la información, que permiten en muchas ocasiones mejorar el acceso y rendimiento del dispositivo. Además, existe una división física y lógica de los discos, la división lógica es en si una división del disco físico para la administración y asignación a determinadas áreas de operación.

Debido a la división del disco en la parte física y lógica, también existen contadores físicos y lógicos. Los contadores de disco lógico miden el rendimiento de los elementos de alto nivel, como conjuntos de volúmenes.

Estos contadores resultan útiles para determinar qué partición provoca la actividad de disco y ayudar a identificar la aplicación o servicio que generan las peticiones. Los contadores de disco físico muestran información acerca de discos individuales, sin tener en cuenta el modo en que son utilizados. Los contadores de disco lógico miden la actividad en las particiones lógicas de los discos, mientras que los contadores de disco físico miden la actividad en todo el disco físico.

Para iniciar el análisis del rendimiento de un dispositivo de almacenamiento secundario, es necesario como regla general identificar primero que no se sufre de problemas de memoria principal, pues los embotellamientos de disco, resultan fáciles de confundir con la actividad de archivo de paginación debido a la escasez de memoria.

Para medir el rendimiento existen unos parámetros que pueden utilizarse, como por ejemplo, el porcentaje de tiempo de disco, utilización del disco, cuellos de botella, etc.

También hay que tomar en cuenta, que existen muchas configuraciones de los discos, y que funcionan como un todo, un ejemplo de esto es un RAID, que es una agrupación de discos y se comportan para su análisis como un objeto único.

2.2.4 Redes

Los objetos de memoria, procesador y disco, son los principales objetos a nivel interno, que afectan el rendimiento del sistema operativo de manera drástica. Pero también existe otro objeto a nivel externo que se comunica, por decirlo así, con otros sistemas operativos, y este tipo de objeto es la red, llamado también, interfaz de red, objeto red, subsistema de red o bien, simplemente red.

El subsistema de red, genera muchas veces embotellamientos de red, principalmente al sistema que está interactuando como servidor, y es necesario comprender algunos conceptos y componentes que son puntos clave para identificar estos problemas.

Los componentes a identificar son: el tipo de carga que generan los clientes, que esta generalmente asociada al tipo de función que tiene el servidor, (servidor de correo, servidor de aplicaciones, servidor de base de datos, etc.), componentes claves de la arquitectura, tipo de red física, tipo de protocolo que utiliza, etc.

Para determinar la carga de cada la red, existen varios contadores que proporcionan la información necesaria para identificar cada uno de los componentes involucrados en la arquitectura de la red.

Si se determina que un subsistema de red tiene problemas de embotellamiento, se pueden realizar diversas medidas para solucionar el problema, entre las cuales se menciona por ejemplo, cambiar controladores por unos más recientes, sustituir los adaptadores por otros mejores o añadir más adaptadores para segmentar la red de manera que se pueda aislar el tráfico de los diversos segmentos. También debe comprobarse el rendimiento general de la red y mejorar los componentes de la capa física, como concentradores y conmutadores, para verificar si la limitación se encuentra en el tendido de la red e incluso distribuir la carga de proceso entre diversos servidores.

Otras soluciones pueden ser modificar el tamaño de los protocolos, para que tengan una mayor cantidad de datos por cada paquete enviado y recibido.

2.2.5 Otras áreas importantes que afectan el rendimiento

Existen diferentes áreas que afectan el rendimiento de un sistema en general, estas áreas, pueden ser divididas en componentes propios del servidor y componentes que hacen uso del servidor.

Entre los componentes propios del servidor estarían lo que es el sistema operativo que trata este trabajo de graduación. Entre los componentes que hacen uso del servidor, están las aplicaciones, como manejadores de base de datos, programas de aplicación, programas de manipulación de datos, en sí estos temas serían para trabajos de graduación adicionales. Actualmente se sabe que existe una tesis sobre el tema de afinamiento de base de datos³, la cual está enfocada al monitoreo y afinamiento de esta área importante. La otra área sobre afinamiento de aplicaciones y recuperación de datos sobre base de datos, son problemas que frecuentemente se tienen en los departamentos de informática, pero pocas veces son tomados en cuenta para realizar procesos de revisión de rendimiento y afinamiento sobre las aplicaciones.

2.3 Objetivo de los componentes de los monitores para la evaluación de los sistemas operativos

Los monitores para evaluar un sistema operativo de manera general, se dividen en tres áreas o componentes principales que son:

- Sistema de búsqueda y recopilación de datos
- Analizador de datos
- Interprete de datos

Estos tratan en sí de minimizar el impacto dentro del sistema operativo en donde se está ejecutando la herramienta de monitoreo y en el sistema operativo

³ La tesis a la que se refiere este comentario, es la que se detalla en la bibliografía a continuación.
Alonzo Lemus, Sergio Rodolfo, Consideraciones para el afinamiento y monitoreo de bases de datos relacionales. Tesis Ing. Ciencias y Sistemas. Universidad de San Carlos de Guatemala, Facultad de Ingeniería, 1999. 226pp.

al cual se desea medir⁴. Pero al ejecutarse generan una carga adicional en los sistemas, principalmente en la computadora en la que se desea medir el rendimiento del sistema operativo, y se busca siempre que esta carga adicional sea tan pequeña que pueda pasar desapercibida en el impacto del rendimiento.

Para el análisis de los sistemas operativos se debe de identificar que funciones realizan cada uno de los componentes de las herramientas de monitoreo, y cómo son determinantes para la evaluación de los sistemas operativos.

Se debe de estar conciente que monitorear las funciones del sistema operativo ayudará a determinar si cambiando determinados parámetros mejorara el rendimiento, o si es necesario adquirir nuevo *hardware*.

2.3.1 Sistema de búsqueda y recopilación de datos

Este sistema comprende el núcleo principal de una herramienta de monitoreo, que es la encargada de solicitar los servicios que proporciona el propio sistema operativo que se está evaluando o bien, de buscar y obtener los datos necesarios para comprender el rendimiento del sistema operativo (interfase de instrumentación).

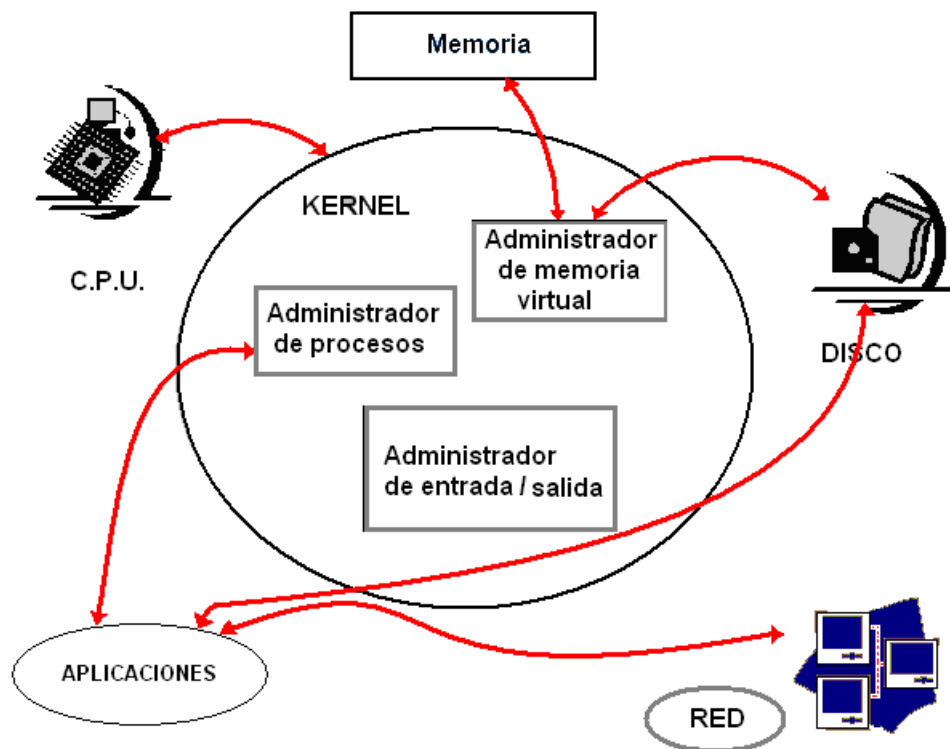
Esta parte de la herramienta de monitoreo incluye los elementos mencionados anteriormente como lo son

⁴ En algunas ocasiones el sistema operativo donde se ejecuta la aplicación de monitoreo y el sistema operativo que se quiere medir o evaluar, es el mismo, aunque es preferible que la aplicación de monitoreo este instalada en otro equipo. No quiere decir necesariamente que sean de diferente tipo de sistema operativo.

- La interfase de la instrumentación
- El selector o filtro de la información
- El procesador de la información a recopilar, etc.

Para lograr efectivamente estas funciones, las herramientas de monitoreo dividen lógicamente al sistema operativo de la siguiente manera.

Figura 2. Estructura del sistema operativo



División lógica de las herramientas de monitoreo de donde el sistema de búsqueda y recopilación de datos puede solicitar información específica de cada una de las áreas que afectan el rendimiento del sistema operativo.

La arquitectura del sistema operativo tal como lo desglosa la Figura 2., es la manera cómo las herramienta de monitoreo perciben lógicamente el sistema

operativo para poder identificar los objetos a monitorear y que son necesarios para medir el rendimiento del sistema operativo.

De esta forma una herramienta de monitoreo puede crear contadores, para medir y evaluar el rendimiento de un objeto, ya sea procesador (CPU), memoria, discos, red como también aplicaciones que están ejecutándose sobre el sistema operativo e identificar el impacto y carga sobre los *objetos* que administran estos recursos, como la interrelación que existe internamente entre un objeto y otro. En consecuencia, un mismo contador es utilizado para describir el comportamiento de varios objetos y la interrelación existente entre ellos.

En sí el administrador de procesos, el administrador de memoria virtual y el manejador de disco, que se ilustran en la parte interior de la Figura 2., son parte fundamental del kernel, que proporcionan parámetros y contadores para medir y evaluar las áreas principales, y en determinado momento la interpretación de varios contadores de de estos administradores, pueden guiar de manera más concreta a las áreas en conflicto del sistema operativo.

Los objetos pueden llevar asociados un número no determinado de contadores que ofrecen información sobre el uso de dispositivos, longitudes de colas, retardos y sobre otros datos utilizados para medir el rendimiento y la congestión interna, contenidos de manera general en el kernel del sistema operativo.

2.3.2 Analizador de datos

El analizador de datos, es la parte del monitor que realiza un preanálisis de los datos, pues muchas veces procesa la información, validando que esté correctamente configurado el sistema, que las medidas sean las correctas, tabulando la información, preparando trazas de información o pares ordenados de datos que serán almacenados para referencias futuras del comportamiento del sistema, para finalmente realizar la función de una etapa previa para el módulo de la interpretación de datos.

2.3.3 Intérprete de datos

La fase de interpretación de datos, consiste en determinar el funcionamiento del sistema que se está monitoreando. Utiliza funciones y reglas para determinar los niveles de aceptación del sistema, incluye en esta parte, algunos sistemas expertos e inteligencia artificial la cual permite con base en una base de datos de información o base de conocimientos, determinar el estado y subsecuentes rendimientos y modificaciones a realizar para tener un rendimiento adecuado de los objetos a monitorear.

En esta etapa se incluye también parte de la presentación de los datos al usuario final, de manera tal que pueda ser interpretada fácilmente, para la toma de decisiones y ajustes del sistema.

Cuando la interpretación de datos se realiza al mismo tiempo de la captura o recopilación de datos, se dice que el sistema de medida es en tiempo real, el cual es de utilidad para tener un control dinámico del rendimiento del sistema.

Por otra parte, si el análisis es diferido, es decir se toman los datos almacenados en la etapa del analizador de datos, toda esta información puede ser importante para la planificación y ajuste del sistema a largo plazo.

Con esta información se puede visualizar gráficamente o con las opciones que permita el monitor, las tendencias de crecimiento de los datos, frecuencias de uso, ciclos de uso, comportamiento del sistema en determinados momentos de carga de datos, ejecución de aplicaciones esporádicas, etc., con el fin de poder configurar o corregir los parámetros del sistema para determinadas tareas o trabajos que son de manera poco frecuentes o de uso no muy frecuente, pero que requieren de un poco más de atención por parte del sistema operativo para lograr un mejor desempeño al momento de entrar en funcionamiento.

2.4 Herramientas gráficas para el monitoreo de los sistemas operativos

Actualmente en el mercado existen muchas utilerías y herramientas que permiten identificar, detectar y resolver conflictos en las principales áreas de rendimiento, estas herramientas son llamadas herramientas de monitoreo, utilerías de afinamiento, monitor de rendimiento, etc., y por su presentación, pueden ser de manera gráfica (que son las que más ampliamente se utilizan por su presentación y facilidad de entendimiento para los problemas e interpretación de funcionamiento) y comandos de líneas que proporcionan datos estadísticos del rendimiento del sistema.

Estas herramientas son parte en algunos casos, de herramientas para afinamiento propias de los sistemas operativos, pero existe *software* de terceros, que se especializan en el monitoreo y afinamiento de los sistemas operativos y se caracterizan por presentar la mayor cantidad de información posible de manera gráfica, que permite a los administradores entender fácilmente la información y determinar las áreas que necesitan especial atención.

La intención de esta sección es presentar las características principales de estas utilerías, y como se describe cada componente de los sistemas operativos para su monitoreo y afinamiento. Posteriormente se describe la representación que utilizan algunas herramientas de monitoreo de terceros y la facilidad para identificar los objetos y áreas con necesidad de afinamiento o mejora en su desempeño.

2.4.1 Opciones de autoconfiguración

Las opciones de auto configuración son opciones propias de los sistemas operativos, que con base en la cantidad de recursos (memoria, espacio en disco, etc.) y la definición u orientación del servidor, preparan y configuran los parámetros de una manera específica.

Algunos de estos tipos de autoconfiguración consiste en indicar que el servidor es para un ambiente multiusuario dedicado al almacenamiento de archivos, programas y textos, en donde las opciones de autoconfiguración, de manera automática maximizan la salida para compartir archivos, asignando la mayor cantidad de memoria al área de caché para disco.

Otra forma es el de maximizar los recursos para una ambiente distribuido, en donde se minimiza el área de memoria para el caché de disco, de manera que se disponga de más memoria para las conexiones y las aplicaciones puedan disponer de una cantidad determinada de memoria.

El inconveniente de estas autoconfiguraciones es que muchas veces realizan más de un cambio a la vez, que afectan otras áreas del sistema, dejando cierta mejora en un área, a un costo de otra área.

2.4.2 Alertas

Las alertas con procesos que se activan cuando un valor registrado por un contador está fuera de los umbrales permitidos. La forma de activación consiste básicamente en notificar o hacer del conocimiento al usuario o administrador el evento registrado.

2.4.3 El visor de sucesos

El visor de sucesos es una utilidad diseñada para hacer un seguimiento de los sucesos grabados en la aplicación, la seguridad y los registros históricos. Permite recopilar información sobre *software*, *hardware* y problemas del sistema, así como hacer un seguimiento de los sucesos de seguridad en el sistema operativo.

Con esta utilidad, se pueden consultar problemas que existen de configuración, problemas ocurridos durante la ejecución de un proceso o bien como resultado de la no disponibilidad de determinado recurso, debido al tipo de origen del problema, se hace una clasificación para almacenarlos y poder describir la información acerca de los mismos.

2.5 Contadores

Los contadores en sí son valores que representan una medida de utilización, sobre determinados objetos (recursos del sistema), que se obtienen con base en una solicitud. Existen muchas formas de poder representar estos valores de manera gráfica.

2.5.1 Velocidad de operación

Este tipo de contador, está orientado para representar una velocidad de operación de parte del objeto, comúnmente se hace referencia al uso de procesador, y detalla un porcentaje de utilización de este, con respecto a los flujos de control de los procesos (*thread*).

Figura 3. Indicador de velocidad de operación



Esta representación del contador gira alrededor de su centro de eje, y lo hace más rápidamente según como sea el valor de utilización del objeto, al centro se representa con un número el valor en porcentaje de utilización del objeto asignado.

2.5.2 Detalle de recursos

Este tipo de contador, detalla de manera exacta una cantidad del mismo tipo de objeto al cual se refiere el contador, así como cuáles están disponibles y cuáles no.

Figura 4. Contador de recursos utilizados y disponibles



Este tipo de contador, detalla que se cuenta con seis objetos del mismo tipo disponibles y cuatro unidades que están en uso o asignadas.

Figura 5. Contador de recursos utilizados



Otra variante de este tipo de contador, es aquella en que se especifica una cantidad discreta de recursos utilizados o bien en espera de que se le asigne un recurso, pero que al llegar a determinado número, debe de hacer referencia a que está con problemas de rendimiento.

Un ejemplo muy útil para ilustrar el uso de este contador es la cantidad de procesos en espera a que el procesador le sea asignado

2.5.3 Factor de utilización y disponibilidad

Estos contadores están orientados a representar de manera continua un rango de posibles valores, que puede mostrar el contador, debido a que cambia de manera muy rápida la utilización del tipo de objeto representado.

Figura 6. Contador de medición continua.



Este tipo de contador a diferencia del anterior, cambian rápidamente los valores, por lo que la proporción asignada y disponible es un valor continuo. Generalmente es utilizado para representar la asignación de la memoria RAM, con un valor máximo disponible y una cantidad asignada.

Figura 7. Contador de medición continua de forma volumétrica.



Otra variante de este tipo de contador es el que tiene forma volumétrica, el cual cumple con los mismos detalles que el anterior, con respecto a representación de valores continuos, pero a diferencia representa una gráfica por cada objeto existente del mismo tipo. Es utilizado para representar el espacio utilizado de los discos de almacenamiento, uno por cada disco detectado.

2.5.4 Contador de procesos

El contador de procesos es un contador que contiene un valor discreto, que puede representar un proceso o un estado específico de un objeto.

Figura 8. Contador de procesos

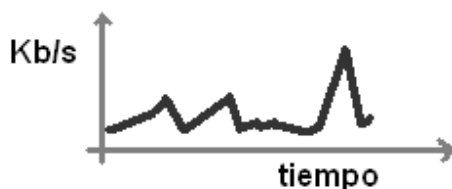


Este contador de procesos, puede identificar adecuadamente valores de tipo, cantidad de servicios, instancias, procesos, etc.

2.5.5 Control de flujo de la información

Este contador es utilizado para representar una cantidad de transferencia de información a través del sistema, esto puede ser entre diferentes objetos, como información transferida entre el disco y la memoria, cantidad de datos escritos de la memoria al disco, etc.

Figura 9. Control de flujo de la información



Los datos a visualizar son valores de velocidad de comunicación, o tamaño de datos intercambiados, en el transcurso de un tiempo determinado.

Generalmente en el eje horizontal se indican las unidades de tiempo, como horas, minutos o segundos y en el eje vertical las unidades de transferencia como Kbytes por segundo o Kbit por segundo, etc.

3 AFINAMIENTO Y RENDIMIENTO, SOLUCIÓN A PROBLEMAS DEL SISTEMA OPERATIVO

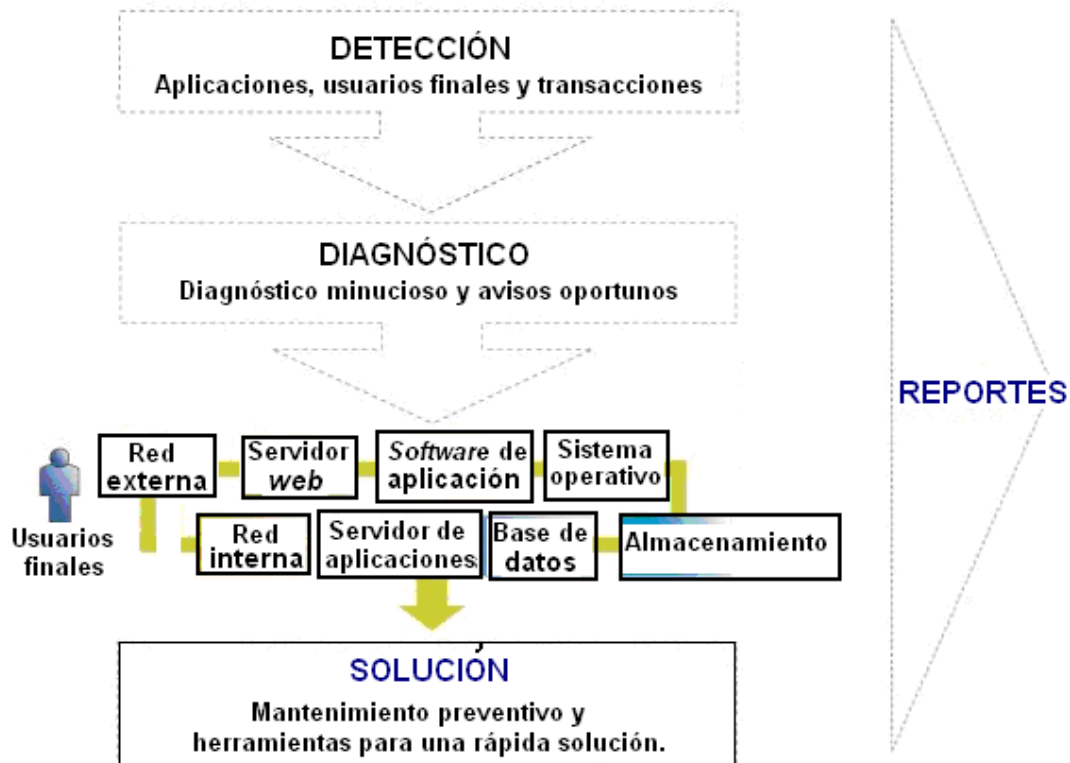
La utilización de monitores, para la solución de problemas de rendimiento del sistema operativo, son todas aquellas herramientas, que ayudan a detectar problemas como

- Cargas de trabajo que no se reparten equitativamente entre los recursos
- Detección de insuficientes recursos del sistema
- Parámetros configurados incorrectamente
- Recursos que funcionan incorrectamente
- Programas que monopolizan los recursos concretos

Otra cualidad de estas herramientas es que también proporcionan información sobre el afinamiento del sistema operativo, es decir que de manera experta proporcionan información adicional sobre que se debe hacer para estos problemas que se detectan, que alternativas de solución son recomendables.

En la figura 10 se muestra el área de detección y diagnóstico, que es parte fundamental en los monitores, y dentro de los cuales se lleva a cabo una serie de procedimientos para proporcionar una solución viable y preventiva para el sistema operativo. En este capítulo se hará énfasis en los aspectos que tienen que evaluar las herramientas de monitores para realizar el diagnóstico y emitir una solución, es decir qué parámetros evaluar y umbrales aceptables para los valores que puede tomar un contador.

Figura 10. Áreas de detección y diagnóstico de los monitores



Representación gráfica de la utilización de los monitores para la detección y resolución de problemas a nivel del sistema operativo.

3.1 Tareas del administrador de sistema

El administrador de sistema, debe de tomar la decisión de cuándo y cómo evaluar el rendimiento del sistema operativo, como también ver como llevar a cabo las tareas necesarias para poder tener un sistema con un desempeño aceptable.

La siguiente tabla, lista las principales tareas que debe de realizar un administrador de sistema, la cual puede disminuir o aumentar según las áreas y tipo de negocio de la empresa.

Tabla II. Tareas de un administrador de sistema

Funcionamiento	Asegurar que el sistema esté inicializado y en funcionamiento. Planificar el mantenimiento preventivo para reiniciar el sistema, (principalmente en algunos sistemas operativos que no cuentan con un eficiente sistema de liberación de recursos). Instalación y actualización del <i>software</i> .
Análisis de fallas	Análisis de fallas con base en bitácoras del sistema y experiencias pasadas.
Planeamiento de capacidad	Conocimiento general a nivel del sistema operativo y planificación de nuevos recursos para demandas futuras.
Afinamiento del sistema	Afinamiento del kernel del sistema operativo, y prioridad de procesos de usuarios para un rendimiento óptimo.
Afinamiento de las aplicaciones	Afinamiento de las aplicaciones para aprovechar al máximo los recursos del sistema.
Planificación de los recursos	Planificación de los procesos, estrategias de seguridad, <i>backups</i> de información, administración y auditoria de los recursos compartidos.

Sistema de redes	Planificación y administración de interconexión de redes, módems e impresoras.
Seguridad	Administración de políticas de seguridad suficientes para impedir ingresos no autorizados como también de mantener una integridad del sistema y privacidad a nivel interno.
Capacitación de los usuarios	Definición de las políticas para ayudar a los usuarios a tener buenos hábitos e instruirlos en el uso del sistema.

Adaptado *Performance Management Guide.*

[http://cepba.upc.es/docs/sgi_doc/SGI_Admin/books/IA_Config/sgi_thml/ch01.tml#id5178992.](http://cepba.upc.es/docs/sgi_doc/SGI_Admin/books/IA_Config/sgi_thml/ch01.tml#id5178992)

En la lista de sugerencias, las tareas de un administrador de sistema son muchas, pero la parte importante para éste trabajo, es el área del afinamiento del sistema operativo y áreas que son parte fundamentales en el rendimiento.

3.2 Objetivos de una evaluación

Los objetivos de una evaluación suelen ser alguna de las siguientes alternativas.

- Comparar alternativas
- Determinar el impacto de una nueva característica, por ejemplo, añadir un disco duro nuevo
- Identificar rendimiento y productividad relativos entre diferentes sistemas.

- Depuración del rendimiento, es decir, identificar los fallos del sistema que hacen que vaya más lento
- Poner unas expectativas sobre el uso del sistema, por ejemplo, cuántas conexiones es capaz de soportar una base de datos simultáneamente, o cuántas peticiones un sitio *web*.
- Afinar el sistema, es decir, hacer que funcione mejor según algún punto de vista (nunca se puede hacer que vaya mejor según *todos* los puntos de vista).

Para el desarrollo de este trabajo el objetivo de la evaluación y medición del sistema operativo es identificar los componentes y áreas para hacer que funcione mejor, como lo indica la última opción, sin embargo, hacer una evaluación no implica que no pueda ser utilizada para las demás características.

3.3 Análisis del afinamiento del sistema operativo

Los procesos que se describen, son de manera general para medir y evaluar el rendimiento de cualquier sistema operativo en general, es decir tomar en cuenta los valores de los parámetros, compararlos con umbrales aceptables según la funcionalidad que tenga con el tipo de carga a trabajar y emitir soluciones al respecto, con el único propósito de que el rendimiento sea lo mejor posible.

Como recomendación se hace la aclaración de que antes de realizar cualquier cambio en los parámetros, se lea qué parámetros deberían ser cambiados y porqué, según el manual de administración de su respectivo sistema operativo.

Generalmente un sistema operativo al ser instalado, está configurado para operar lo más rápido posible bajo circunstancias comunes. Como sea, siempre es posible ajustar ciertos parámetros y valores del sistema operativo para mejorar el rendimiento u optimizarlo, para que sea mucho más efectivo bajo ciertas condiciones de trabajo que se desea, por ejemplo, para que el rendimiento sea mucho más eficiente con el acceso al disco; o requerimientos específicos sobre una aplicación de *software* instalada, para que el *software* sea más eficiente.

3.3.1 El arte del afinamiento del sistema operativo

Existen varias consideraciones, que se deben mencionar sobre el afinamiento del sistema operativo y que se deben de tener en cuenta durante el proceso; el afinamiento es un arte más que una ciencia, pues muchas veces es necesario repetir varias veces un mismo proceso para el afinamiento. También se dará cuenta que es necesario hacer un monitoreo extensivo y pruebas a fondo, repetidamente, para afinar el sistema.

3.3.2 Afinamiento y rendimiento del sistema

El afinamiento del sistema, es la respuesta a preguntas fundamentales como: ¿Cómo puedo obtener un mejor rendimiento con el *software* y *hardware* actual, sin necesidad de agregar más?, Algunos problemas únicamente se solucionan adquiriendo más *hardware* (que pueden ser más grandes de capacidad y rápidos, como los discos), pero muchos problemas pueden ser resueltos haciendo un mejor uso o configuración de los recursos existentes, a esto es a lo que se llama afinamiento.

Para poder realizar el desempeño del afinamiento, es necesario entender que el arte de afinar es en sí un asunto de administración de recursos y configurar correctamente los parámetros del sistema. Afinar el sistema con una determinada carga de trabajo, para un eficiente uso de recursos consiste en realizar los siguientes pasos siguiendo un lineamiento de tareas y metodologías que a continuación se detallan.

- a. Identificar los objetivos del afinamiento
 - Determinar como se medirán los resultados
 - Cuantificar y priorizar los objetivos
- b. Identificar los recursos críticos que limitan el rendimiento del sistema
- c. Minimizar los requerimientos de los recursos críticos de la carga de trabajo del sistema
 - Utilización de recursos más apropiados
 - Reducción de requerimientos de recursos críticos para programas individuales o funciones del sistema
 - Estructurar el uso paralelo de los recursos
- d. Modificar la asignación de recursos para reflejar las prioridades
 - Cambiar la prioridad o los recursos limitados
 - Cambiar la configuración de parámetros del administrador de recursos del sistema
 -
- e. Repetir los paso de b a c hasta que los objetivos sean cumplidos (o los recursos estén saturados)
- f. Asignar recursos adicionales al sistema de ser necesario.

3.3.2.1 Objetivos al momento de realizar un afinamiento

Generalmente se puede mejorar el rendimiento del sistema con la misma carga de trabajo, pero se deben de tomar en cuenta factores al momento de realizar el afinamiento, por lo que es necesario definir objetivos al realizar un afinamiento del sistema operativo. Cuando se ha determinado la carga de trabajo del sistema operativo, se debe de seleccionar el o los criterios del afinamiento y el conjunto de objetivos del afinamiento basados en los criterios seleccionados.

Los principales criterios para el afinamiento están basados en dos puntos que son:

- a) Tiempo de respuesta, que es el tiempo transcurrido desde que se ha solicitado una determinada tarea y la respuesta a la solicitud es presentada. Ej. ¿Cuánto tiempo le toma a la base de datos realizar una consulta?
- b) Rendimiento del sistema (*throughput*), se refiere a la medida de qué tanto trabajo se realiza por unidad de periodo de tiempo. Ej. ¿Qué cantidad de transacciones realiza la base de datos por minuto?

La relación entre estos dos criterios puede ser compleja, en muchos casos se debe decidir entre uno u otro criterio, es decir tener un mayor rendimiento a un costo de un mayor tiempo de respuesta; o bien tener un menor tiempo de respuesta a un costo menor de rendimiento. En algunas otras situaciones un pequeño cambio en los parámetros del sistema operativo, puede mejorar significativamente ambos criterios de afinamiento.

Un rendimiento aceptable es basado en un rendimiento razonable, combinado con un tiempo de respuesta razonable.

Al definir los objetivos, estos tienen la característica que pueden ser medidos de manera cuantificable, además en algunos casos, que pueden ser subjetivos, tal como un tiempo de respuesta satisfactorio.

3.3.2.2 Identificación de los recursos críticos

El rendimiento del sistema operativo, está determinado por la disponibilidad y velocidad de uno o dos recursos críticos del sistema. El analista, debe de identificar estos recursos correctamente o correr el riesgo de caer en un proceso de prueba y error en afinar recursos al azar.

El sistema operativo como tal, tiene una división lógica y física de los recursos. Los recursos físicos, son generalmente fácil de identificar, pues muchas herramientas de afinamiento cuentan con medios para evaluar la utilización de estos recursos físicos. Los recursos físicos que más afectan el rendimiento son:

- Ciclo de procesador
- Memoria
- Bus de Lectura/Escritura de Disco (I/O)
- Adaptadores varios
- Brazos de discos
- Espacio de discos
- Acceso a red.

Los recursos lógicos son más difíciles de identificar. Estos son generalmente particiones de los recursos físicos, realizados por algunos programas. El objetivo de una partición lógica, es generalmente con el fin de compartir y administrar los recursos físicos.

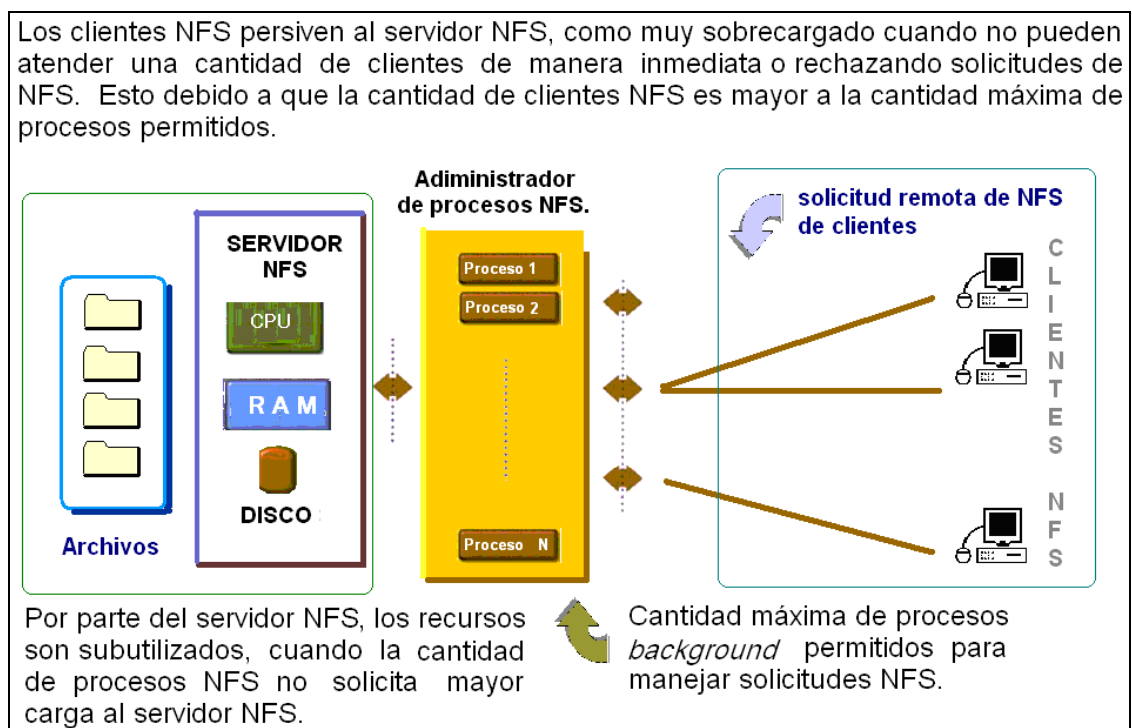
Tabla III. Clasificación de los recursos físicos y lógicos

RECURSO REAL/FÍSICO	RECURSOS LÓGICOS O PARTICIONES
Procesador	Tiempo de asignación del procesador (<i>quantum</i>)
Memoria	Paginación <i>Stacks</i> <i>Buffer</i> Colas Tablas Bloqueos y semáforos
Disco	Volúmenes lógicos Sistema de archivos (<i>file system</i>) Archivos Particiones
Red	Sesiones Paquetes Canales Puertos

Hay que tener en cuenta lo importante que son los recursos lógicos como los recursos físicos. Un proceso puede ser bloqueado por la falta de recursos lógicos, como por recursos físicos, y expandir los recursos físicos no implica que se añadirán los recursos lógicos necesarios, por lo que es necesario verificar la expansión de los recursos lógicos necesarios.

Como ejemplo de la falta de definición de recursos lógicos, la siguiente figura, cuenta con un servidor de archivos que administra las solicitudes de bloques de información con un sistema NFS, y se requiere de un proceso *background* por cada solicitud remota de NFS de los clientes.

Figura 11. Solicitud de servicio NFS



Problemas de rendimiento debido a la configuración de capacidad máxima de procesos permitidos, (procesos *background*), son identificados como recursos lógicos que no pueden corregirse con aumento de memoria, procesador, disco, etc.

Cada uno de los procesos es un recurso lógico y pueden trabajar simultáneamente y cuando se llega al límite máximo de la cantidad de procesos *background* permitidos, limita la capacidad de atención del sistema NFS, dando la impresión de que el disco, procesador y puertos de comunicación, están siendo ligeramente utilizados o sin carga, cuando en realidad hay una limitante de la cantidad de solicitudes de NFS para atender. En un momento dado, se puede tener la impresión de que el sistema en general está siendo subutilizado y los clientes tienen la impresión de que el sistema está lento, cuando en realidad, se tiene definido una pequeña cantidad de procesos *background* de NFS, que limitan el resto de los recursos.

Los procesos *background* necesitan ciclos de procesador y memoria, pero agregar más memoria física o un procesador más o más rápido no es la solución. La solución es configurar adecuadamente la cantidad de procesos *background* que manejan las solicitudes al sistema NFS. En otras palabras es necesario agregar más recursos lógicos, que administran las solicitudes al sistema NFS.

3.3.2.3 Minimizando los requerimientos de los recursos críticos

La reducción de los requerimientos de los recursos críticos, es el proceso de administración de los recursos para lograr un aprovechamiento más efectivo o mejor aprovechado. Esto lo se puede realizar a través de tres formas.

- **Uso apropiado de los recursos:** es la elección de los recursos con base en los objetivos que se desean lograr. Por ejemplo, la utilización apropiada de recursos, se puede mencionar en la implementación de una nueva aplicación, cuando se decide aumentar el uso de la memoria, para liberar el uso del procesador. Otra configuración similar, es decidir, donde se localizaran los archivos, ya sea localmente o en un servidor específico.
- **Reducción de los requerimientos de los recursos críticos:** es la revisión de las aplicaciones locales, o desarrolladas de manera interna, para mejorar el desempeño o bien reducir funciones innecesarias. Los procesos de poca importancia o poca carga de información, pueden ser reubicados a otros servidores, o ser ejecutados en horarios, etc.
- **Estructuración de los recursos para ser usados paralelamente:** es la división de los recursos, para que las tareas que requieren de múltiples recursos puedan ser ejecutadas satisfactoriamente, la división de los recursos se realiza para que estos sean utilizados paralelamente. Por ejemplo, si una aplicación debe de leer dos archivos secuencialmente, entonces si se agrega otro disco al sistema con un manejador adicional o diferente, podemos mejorar el acceso de lectura/escritura de los discos y aprovechar esta división, para que los archivos sean leídos paralelamente, uno en cada disco.

3.3.2.4 Modificación de la asignación de los recursos para reflejar la prioridad

El sistema operativo provee de muchas formas para modificar la prioridad de los recursos disponibles, algunos son configurados a bajo nivel, como la velocidad de los discos, otros pueden ser configurados a nivel de usuarios individuales o grupos, de manera que puedan ser asignados de manera de prioridad de las tareas a realizar, por ejemplo, la asignación del procesador.

3.3.2.5 Repetición de los pasos del afinamiento del rendimiento

Desafortunadamente, en el área de afinamiento del sistema operativo, siempre que se realice una mejora, aparecerá otra área que necesita ser afinada o mejorada, este es un ciclo continuo de nunca acabar, por lo cual, como se menciona al inicio de esta sesión, es importante definir el objetivo u objetivos al inicio de afinamiento, pues muchas veces, se resulta afinando áreas que son secundarias.

Además, se debe de tomar en cuenta que al realizar un afinamiento, se incurre en sobrecargar otros objetos, lo importante es tener siempre en mente los objetivos del afinamiento.

3.4 Evaluación del rendimiento y afinamiento de la memoria

La memoria es el recurso del sistema operativo que constantemente está lleno a su capacidad máxima. Incluso cuando aparentemente se encuentre en ejecución una única aplicación y que no consuma toda la memoria. Esto es debido a que el sistema operativo mantiene en memoria páginas de datos, que son referencias de las últimas aplicaciones ejecutadas, datos utilizados por las aplicaciones, de manera que si un proceso requiere nuevamente utilizar la información que tiene presente, pueda minimizar el acceso a disco (lectura y escritura). Este proceso por mejorar tiene sus costos que implican que cuando existe demanda de memoria, es necesario satisfacerla y muchas veces implica liberar memoria para asignarla a quien la solicita, pero ¿qué bloques de memoria liberar de manera que sea el rendimiento más adecuado y cómo hacerlo?

3.4.1 Problemas de la memoria

Los problemas de memoria pueden presentarse con otros síntomas de falla, como una sobrecarga en el procesador o un funcionamiento perezoso del disco. El mejor indicador de un embotellamiento de memoria es una tasa alta y constante de errores de página de *hardware* (más de cinco por segundo, por ejemplo). Estos errores se producen cuando un programa no encuentra en la memoria física los datos que necesita y tiene que recuperarlos del disco.

3.4.1.1 Pérdida de la memoria

El problema de la pérdida de memoria por las aplicaciones o retención de la memoria no utilizada por las aplicaciones, es debido a la liberación incorrecta de la memoria por algunas aplicaciones. Esto ocurre cuando la memoria es asignada y se detecta un aumento uniforme en la asignación de memoria a una misma aplicación, el cual sugiere que la aplicación no libera la memoria asignada previamente de manera correcta. Sin embargo, se debe de tener en cuenta el tipo de aplicación y el comportamiento de ésta, previo a emitir este tipo de opinión. La cantidad de memoria asignada a una aplicación recibe el nombre de espacio de trabajo.

3.4.1.2 Paginación

La paginación dentro de la administración de memoria, es otro factor que puede causar deterioro en el rendimiento del sistema operativo, pues generalmente existen uno o más archivos para paginación, los cuales al aumentar su tamaño de manera automática más allá del tamaño ideal o del tamaño predeterminado, para satisfacer las demandas de memoria por el sistema operativo, este empieza a deteriorar el rendimiento, especialmente si el archivo crece un una partición lenta o en un disco fragmentado.

La paginación consiste en mover partes (o secciones) de memoria de un proceso a disco (en unidades llamadas páginas). Si un proceso necesita acceder y tener más páginas que no están actualmente en la memoria (y que están escritas en el disco), ocurre lo que se llama *page fault*.⁷ Lo cual implica leer las páginas y cargarlas a la memoria.

De hecho el *swap* y la paginación, son procesos normales del sistema operativo, Pero el problema ocurre cuando el sistema operativo necesita realizar *swap* y paginación excesiva,

3.4.1.3 *Swaping* (alternar o intercambiar)

Swaping o swap significa que la información del proceso completo es escrita a disco, cuando por cualquier motivo deja el control de procesador. La próxima vez que el sistema asigne el procesador al proceso, necesita leer completamente el proceso del disco.

3.4.2 Parámetros para medir y evaluar el uso de la memoria

Los parámetros que pueden ayudar a medir los problemas en la utilización y rendimiento de la memoria son:

⁷ *Page faults* es cuando el sistema operativo necesita información y no se encuentra en la memoria por lo cual es necesario leerlo y cargarlo a la memoria, generalmente se carga un tamaño fijo de datos a la memoria en bloques denominas páginas de memoria y generalmente son de un tamaño de 4Kbyte.

3.4.2.1 Páginas por segundo

Muestra el número de páginas solicitadas que no se encuentran disponibles inmediatamente y que deben leerse del disco, o bien que han tenido que escribirse al disco para liberar espacio destinado a otras páginas. Si este número es elevado cuando el sistema se encuentra bajo una carga normal, se debería pensar en aumentar la memoria física.

Si este parámetro de las páginas por segundo aumenta pero el contador de memoria que nos indica la cantidad de bytes disponibles se reduce al límite de memoria disponible y los discos que contienen los archivos de paginación están ocupados (indicado por los parámetros «porcentaje de tiempo de disco», «bytes de disco/segundo» y «longitud promedio de cola de disco»), es un indicador de embotellamiento de memoria. Si el parámetro «memoria: bytes disponibles» no disminuye, es posible que no sea este el problema. En ese caso, se debe de comprobar si hay alguna aplicación que lleve a cabo muchas lecturas y escrituras del disco y que los datos no se encuentran en la memoria caché.

Cuando se produce la presión de memoria excesiva, las peticiones de memoria provocan fallos de página de *hardware* a un ritmo alarmante y el sistema operativo dedica la mayoría del tiempo a retirar y colocar datos en la memoria virtual, en lugar de ejecutar aplicaciones. La presión de memoria excesiva agota rápidamente los recursos del sistema y suele aumentar los tiempos de respuesta. Si una aplicación deja de responder pero el indicador de actividad del disco duro sigue parpadeando, es probable que el equipo esté sufriendo este problema.

Un valor constante mayor a 5 páginas por segundo de memoria puede ser considerado como un problema de memoria.

3.4.2.2 Bytes comprometidos

Indica la cantidad de memoria virtual que ha comprometido el sistema en la memoria física para almacenamiento o en el archivo de paginación. Si el número de bytes comprometidos es mayor que la cantidad de memoria física, es necesaria más memoria física.

La memoria comprometida no debería de ser de más del 75 por ciento del total de la memoria física.

3.4.2.3 Memoria virtual activa

La memoria virtual activa es la cantidad de memoria virtual utilizada por las aplicaciones y procesos, implica de manera explícita uso adicional del procesador y acceso al disco, para almacenar las páginas de memoria virtual.

Este parámetro de la cantidad de memoria virtual activa, es un promedio de la cantidad de páginas de memoria virtual que han sido utilizadas por el sistema operativo, es ideal para poder determinar una cantidad apropiada de memoria física (RAM) para un sistema. Idealmente se debe de tomar el máximo valor reportado en un determinado periodo y convertirlo o expresarlos en *bytes*, pues muchos sistemas lo representan en cantidad de páginas o bloques de memoria de cierta cantidad de *bytes*.

Idealmente el valor de bytes de la memoria virtual debería de estar por debajo del total de memoria RAM, de lo contrario cierta cantidad de paginación de memoria virtual existiría en el sistema. Qué tanta paginación ocurra, dependerá de la diferencia entre estos dos valores (la cantidad de memoria RAM y la cantidad de memoria virtual activa).

Si se tiene más memoria virtual que la memoria física real (RAM), esto podría causar una paginación excesiva la cual provocaría como resultado un retardo por falta de memoria.

Si el promedio de la memoria virtual activa, es inferior a la memoria RAM, entonces puede ser que la cantidad de páginas para el espacio de paginación pueda ser la causante de que la memoria RAM se llene. En este caso se debe de verificar los parámetros que definen la cantidad de páginas para paginación de archivos y las páginas para datos. En otras palabras, qué cantidad de memoria es utilizada para datos y qué cantidad es utilizada para archivos y programas.

3.4.2.4 Porcentaje de uso del archivo de paginación

Muestra el porcentaje en uso del archivo de paginación que ha utilizado el sistema operativo. Si este valor alcanza o supera el 80 por ciento, se debe de pensar en aumentar el archivo de paginación. Generalmente el tamaño óptimo del archivo de paginación, debe de ser de un tamaño igual a la cantidad de memoria real existente más el tamaño de encabezado de los archivos, que puede ser de aproximadamente 12 Kbytes.

3.4.2.5 Bytes de memoria no paginable

Indica la cantidad de memoria no paginable, que es un área de la memoria del sistema donde los componentes adquieren espacio a medida que realizan tareas asignadas. Si este valor muestra un crecimiento constante pero no se aprecia el correspondiente incremento en la actividad del servidor, es posible que algún proceso en ejecución tenga pérdidas de memoria. Estas pérdidas se producen cuando un defecto de programación impide que un programa libere la memoria que ya no necesita. Con el tiempo, estas pérdidas de memoria pueden provocar el fallo del sistema, ya que toda la memoria disponible, tanto física como la del archivo de paginación, está toda asignada. Cuando la cantidad de memoria no paginable es demasiado alta, es necesario adicionar más memoria física, para satisfacer esta demanda.

3.4.2.6 Memoria física disponible

La cantidad de memoria física disponible es un parámetro que especifica la cantidad de páginas de memoria mínima que el VMM (*Virtual Memory Manager*) debe de mantener en la lista de páginas de memoria. Cuando la cantidad de páginas de memoria física llega a este punto es cuando se tiene la necesidad de recurrir a espacios de memoria virtual, en otras palabras cuando este valor de la cantidad de páginas de memoria física es pequeño, el sistema estará constantemente leyendo páginas de memoria y descargando páginas de memoria al disco, este proceso es conocido como *thrashing*.

3.4.3 Soluciones a los problemas del uso de la memoria

Cuando se sufre de problemas de rendimiento por memoria, deben de observarse la paginación. Si el sistema se encuentra paginando frecuentemente, puede disponer de poca memoria en la máquina. Algo de paginación es bueno, puesto que ayuda a expandir la memoria de algún modo, pero un exceso de paginación es una carga para el rendimiento del sistema.

3.4.3.1 Administrador de la memoria

El sistema operativo, para administrar la memoria, utiliza un procedimiento llamado administrador de memoria virtual (*Virtual Memory Manager* VMM), que tiene como función la asignación de páginas de memoria real y la resolución de las referencias de las aplicaciones o programas a páginas de memoria virtual que actualmente no están en memoria real o que aún no existen.

El administrador de memoria virtual tiene dos objetivos, aunque a veces opuestos:

- Minimizar los costos de sobrecarga del tiempo de procesamiento y utilización del disco para el uso de la memoria virtual.
- Minimizar los costos de tiempo de respuesta de *page faults*.
-

3.4.3.2 Tamaño de paginación y *swap*

La paginación se utiliza para liberar memoria para otros usos moviendo bloques de datos y código de un tamaño fijo desde la RAM hasta el disco en unidades denominadas páginas. Un tamaño ideal para los archivos de paginación como para *swap*, debería de ser del tamaño de la memoria real más el tamaño de los encabezados de los archivos, que es en promedio de 12 *Kbytes*.

Para determinar si el sistema está paginando en exceso, debe de tomarse los datos de los parámetros <<memoria: páginas por segundo>>, que se detalló previamente y <<disco: promedio en segundos por transferencia>>, que muestra la media de transferencias al disco en segundos. Si estos dos parámetros al ser multiplicados excede de 0.1 (esto es 10 por ciento), significa que la paginación está ocupando más del 10 por ciento del tiempo de los acceso al disco, la cual es considerada como un valor alto de paginación.

3.4.3.3 Programas con uso de memoria intensivo

Si se ha identificado programas que hacen uso intensivo, que no se pueden modificar o mejorar, una solución es que deberían de ejecutarse cuando la carga del sistema sea pequeña o en computadoras de alto rendimiento.

3.4.3.4 Páginas de memoria de archivos y datos

La memoria real, es dividida en páginas de memoria. Estas páginas de memoria son utilizadas para almacenar datos de archivos o datos de cálculo o texto de información, que fueron utilizados recientemente por el sistema operativo o bien que están siendo utilizados.

Existen en la configuración del sistema operativo parámetros que definen un cantidad mínima y máxima de páginas de memoria que pueden ser utilizadas para archivos y datos de operación. Las páginas que recientemente fueron utilizadas por el sistema operativo, permanecen en la memoria por un tiempo por si son requeridas nuevamente, de manera tal que pueda reducir operaciones al disco de lectura/escritura. Estos parámetros también son ideales para orientar el uso del servidor, es decir si se requiere tener un servidor para archivos, es ideal incrementar el límite superior de la cantidad de páginas para archivo, en caso de requerirse el servidor para procesamiento de datos y cálculos con pocas referencias a archivos, es ideal reducir este parámetro.

Idealmente el tamaño de las páginas de memoria debería de ser del tamaño de los bloques de disco o múltiplos de este tamaño.

Tabla IV. Parámetros para la evaluación y rendimiento de la memoria

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIA
Páginas por segundo	Número de páginas solicitadas que no se encuentran disponibles inmediatamente en memoria y deben de ser solicitas al disco	Máximo 5 unidades	Un elevado número de páginas indican una alta paginación
Bytes comprometidos	Cantidad de memoria comprometida por el sistema operativo	Máximo 75% del total de la memoria	Si se sobrepasa este límite es necesario considerar aumentar la memoria. No es recomendable que se comprometa más de éste porcentaje
Memoria virtual activa.	Cantidad total de memoria utilizada por el sistema operativo.	Igual a la cantidad de memoria real.	Si el total de memoria virtual activa es mayor a la cantidad de memoria real, entonces existirá paginación en el sistema. Factor ideal para determinar la cantidad de la memoria del sistema.

Continuación 2/3

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIA
Porcentaje de uso del archivo de paginación	Indica el porcentaje de utilización de los archivos para paginación	Menor posible	Un alto porcentaje es un indicador de más memoria requerida. No es recomendable que exista paginación excesiva
Páginas de memoria física disponible	Cantidad mínima de páginas de memoria física disponible que el sistema operativo debe de poseer	Mínimo 15 % del total de la memoria	Cuando el total de la memoria disponible llega a este valor, el sistema operativo realiza paginación para tener memoria disponible para futuros requerimientos. Por debajo de este porcentaje, se debe aumentar la cantidad de la memoria física.
Tamaño del archivo de paginación.	Indica el tamaño total de los archivos utilizados para paginación.	El total debe de ser del tamaño de la memoria virtual más el tamaño del encabezado de cada archivo.	Crear un archivo por cada unidad de disco disponible. El total de los archivos creados para paginación debe de ser como máximo del tamaño de la memoria física

Continuación 3/3

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIA
Tamaño de los archivos para <i>swap</i>	Indica el tamaño total de los archivos utilizados para <i>swap</i>	El total debe de ser del tamaño de la memoria virtual más el tamaño del encabezado de cada archivo creado para <i>swap</i>	Utilizar dispositivos de almacenamiento más rápidos disponibles. Un factor alto de <i>swap</i> , indica que es necesario ampliar el tamaño de la memoria física
Porcentaje de paginación	Total del tiempo utilizado para paginación	Menos del 10% de utilización	Aumentar la memoria para evitar mucha paginación

3.5 Evaluación del rendimiento y afinamiento del procesador o CPU

El procesador es uno de los componentes más rápidos del sistema, y es casi imposible que un único programa mantenga ocupado al procesador en un 100% por un periodo considerable (si mucho por un instante, y tener ocupado al sistema, se entiende que el procesador esta 0% disponible y 0% en espera), pero aún cuando se trata de un ambiente multiusuario o multiprogramación, también es raro que el procesador trabaje a un 100% por un periodo considerable. En estos casos puede tratarse de un programa que se encuentra en un ciclo infinito de ejecución, aún cuando el programa sea demasiado tardado, debe de identificarse y revisarse antes de ser eliminado el proceso que está sobrecargando el uso del procesador.

Los principales componentes de un procesador son los procesos y los *threads* creados. Cada proceso puede crear uno o más *thread*, y cada *thread* compite por el uso del procesador, con otros *threads*, de otros procesos. Un proceso es una actividad que ha iniciado en el sistema a partir de un comando, o un programa desde el *shell* o *prompt*, u otro proceso. A partir de aquí, se pueden analizar una serie de etapas, estados y atributos de cada uno de estos componentes.

Para monitorear el uso del procesador, existe una diversidad de comandos, como aplicaciones que proporcionan información gráfica del rendimiento del uso del procesador. Esto aplica generalmente para el monitoreo de cualquier objeto o recurso del sistema operativo a medir. Pero lo importante de esto, es que se pueda interpretar y saber que acciones se deben tomar para mejorar el rendimiento según los objetivos del proceso del afinamiento.

3.5.1 Problemas con el procesador

Los problemas con el procesador, generalmente embotellamientos, se producen únicamente cuando el procesador está tan atareado que no puede responder a las peticiones. Entre los síntomas de esta situación se encuentran los elevados índices de actividad del procesador, grandes colas de forma frecuente y un mal tiempo de respuesta de las aplicaciones. Las aplicaciones y controladores dependientes del procesador y las interrupciones constantes, provocadas por componentes mal diseñados de los subsistemas de red o de disco, son causas habituales de problemas de rendimiento del procesador.

Un procesador únicamente puede ejecutar un proceso a la vez en un instante determinado. Si está sobrecargado, los procesos tienen que esperar en lugar de ejecutar las tareas asignadas. El problema principal, es que no se puede modificar o aumentar la velocidad, lo que se puede hacer es cambiar el procesador por uno más rápido o adicionar otro (si lo permite la arquitectura del *hardware*), pero esta es una opción muchas veces difícil por el factor económico. Otra solución es monitorear el uso del procesador y tratar de distribuirlo para que sea más eficiente.

3.5.1.1 Programas mal diseñados

Un procesador balanceado en un ambiente multiusuario, indicará en actividades de usuario entre un 50-55%, actividades del sistema entre un 30-35% y actividades de espera por lectura/escritura de disco entre 10 y 15% de tiempo de uso del procesador.

Cuando un programa mal diseñado, ha caído en un ciclo infinito, y se apropia del uso del procesador, las actividades del usuario aumentan significativamente entre un 70 a 90% de asignación del procesador y las actividades del sistema entre un 10 y 30% aproximadamente, sin estar en espera de lectura/escritura del disco (esto es el tiempo de espera del disco es 0%), además que esta actividad es prolongada por un tiempo considerable.

Cuando no existen procesos o *threads* en espera, se identifica en el uso como *idle* (disponible), y a este parámetro es asignado todo el tiempo del procesador. Cuando se da esta característica, indica una subutilización del procesador, lo cual de ser recurrente, se puede modificar algunos parámetros para realizar más actividades con el procesador y así, liberar o ayudar otro recurso que puede estar muy atareado.

3.5.1.2 Cambios de contextos

Los cambios de contextos (*context switch*) se refieren a la cantidad de intercambios de *threads* en el procesador en un intervalo de tiempo. El procesador como recurso es dividido en unidades lógicas de tiempo de asignación a los procesos, generalmente de 10 milisegundos cada uno⁸. De manera tal que un *thread* es asignado al procesador, este se ejecuta hasta que el tiempo expira o deja el procesador porque se agota el tiempo de manera voluntaria o de manera forzada.

⁸ El valor aquí presentado, es un valor promedio en un sistema operativo como AIX, el cual en otros sistemas puede variar la unidad o tamaño.

Cuando otro *thread* toma el control de procesador, el contexto o ambiente de trabajo del *thread* previo, debe de ser guardado para que el nuevo sea leído y tome el control. Cada cambio de *thread* que hace uso del procesador, es bastante rápido y no requiere de una cantidad grande de recursos. Por lo tanto, cualquier incremento significativo en el cambio de contextos, indica un potencial problema de lectura/escritura del disco o de problemas de tráfico de la red, que afectan los procesos en el procesador, cada *thread* se ve obligado a dejar el control, por espera de un recurso ocupado o no disponible.

3.5.2 Parámetros para evaluar y medir el uso del procesador

Cuando se haga un seguimiento del uso del procesador, habrá que considerar el rol de la computadora y el trabajo que se encuentra realizando. Los valores de procesador altos pueden significar que la máquina se está encargando de manejar la sobrecarga del sistema de una manera eficiente o por el contrario, puede estar constantemente luchando para poder continuar.

3.5.2.1 Interrupciones por segundo

Este contador puede utilizarse para averiguar si la actividad de las interrupciones es la causa de un embotellamiento del procesador. Si se observa un incremento dramático de este contador sin un incremento correspondiente en la actividad del sistema, probablemente se trate de un problema de *hardware*. Para resolver el problema, se necesitará encontrar los dispositivos que estén causando las interrupciones, y ver en las especificaciones del fabricante el valor del umbral del procesador que se considere aceptable.

3.5.2.2 Porcentaje de tiempo de interrupción

Este contador visualiza el porcentaje de tiempo que el procesador pasa enviando y atendiendo interrupciones de *hardware* durante el intervalo de muestreo. Este valor da una indicación general de las actividades de los dispositivos que generan las interrupciones, como por ejemplo, unidades de disco, adaptadores de redes y otros dispositivos periféricos.

Estos dispositivos interrumpen al procesador cuando requieren su atención o completan una tarea. Debe observarse si se producen incrementos dramáticos en el valor sin que se produzca un incremento correspondiente en la actividad del sistema.

3.5.2.3 Tiempo de uso del procesador

Un uso óptimo de procesador sería que éste trabajara a un 100%. Esto sería cierto en el caso de un sistema monousuario que no necesita compartir el procesador. Generalmente, si el tiempo de uso del procesador por parte de las aplicaciones del usuario (*us*) más el tiempo de procesador por el sistema (*sy*) es por debajo del 90%, en un sistema monousuario, éste no sería considerado un problema.

$us + sy \leq 90\%$ tiempo de uso del procesador \Rightarrow rendimiento aceptable.

Pero para un sistema multiusuario, tenemos que

$us + sy \geq 80\%$ tiempo de uso de procesador \Rightarrow procesos en cola.

Lo que significa que los procesos pueden estar un tiempo en espera en la cola de procesos. El tiempo de respuesta y el rendimiento en este caso puede adolecer problemas por uso del procesador.

Un valor elevado en el porcentaje de tiempo del procesador no siempre implica la existencia de un embotellamiento. Si el procesador atiende todas las peticiones sin aumentar las colas de trabajo del servidor ni la longitud de cola del procesador, esto es que atiende los procesos y está trabajando sin embotellamiento. El embotellamiento se produce cuando el valor del parámetro «sistema: longitud de la cola del procesador» aumenta, y el valor del parámetro «procesador: % de tiempo de procesador» se mantiene alto y la memoria, la interfaz de red y los discos no presentan ningún embotellamiento.

3.5.2.4 Colas de procesos

Cada procesador tiene una cola de proceso, la cual es una lista ordenada por prioridad de los *threads*, que están listos para ser atendidos por el procesador. El valor dado en el parámetro de cola de procesos por las herramientas de monitoreo (cola de *threads* en espera) es en sí la suma de *threads* listos para ejecución en espera de asignación del procesador. Si se divide este número, entre la cantidad de procesadores que cuenta el servidor, nos proporciona un número promedio de los *threads* que pueden ser ejecutados o atendidos por el procesador. Si este valor es mayor que uno (1), significa que los *threads* deben de esperar su turno para que se les asigne el procesador. Entre mayor es el número promedio, mayor va a ser el retardo de atención del procesador, lo que significa una sobrecarga.

Otro parámetro relacionado con la cola de procesos, es el número de *threads* que están en la cola, pero no están listos para ejecución, estos *threads* son los que están en espera a que se libere un recurso o en espera de que se complete un proceso de lectura/escritura por parte del disco.

Las estadísticas referentes a la cola de espera de procesos, de los procesos listos para ejecución, son valores promedios que en un sistema no-SMP (*no-simetric multiprocessing* o sin multiprocesamiento simétrico) este valor debería de ser menor de 5 unidades⁹.

Proceso cola > 5 => problemas con asignación del procesador.

Pero en un sistema SMP (multiprocesamiento simétrico), este valor debería de ser menos de:

$$5 \times (N_{\text{total}} - N_{\text{bind}})$$

Donde N_{total} es el número total de procesadores y N_{bind} es el número de procesadores que han sido asignados a un determinado proceso.

⁹ En algunos sistemas operativos restringen este valor a no mayor de 3 unidades, pero principalmente depende de la orientación del sistema y de los objetivos de afinamiento.

Si este número se incrementa rápidamente, es necesario examinar las aplicaciones. Sin embargo, el sistema debería de responder adecuadamente con 10 a 15 *threads* en las respectivas colas (haciendo recordatorio que se define una cola por procesador), dependiendo de la tarea del *thread* y la cantidad de tiempo de ejecución.

3.5.2.5 Medición del uso del procesador por la ejecución de las aplicaciones

Es importante determinar la cantidad de tiempo asignado del procesador a cada una de las aplicaciones. Esta medida es llamada tiempo real y es el tiempo transcurrido desde el inicio hasta la finalización de la aplicación. En el tiempo real, se determina el tiempo usado por la aplicación. El tiempo es dividido entre el usuario y el sistema. Los valores del usuario es el tiempo usado por las aplicaciones en sí y cualquier librería o subrutinas que este llama. El valor del sistema es el tiempo usado para llamadas del sistema, invocadas por la aplicación de manera directa o indirecta.

La suma del tiempo del usuario más el tiempo del sistema, es el costo directo total de procesador para la ejecución de las aplicaciones. Este costo no incluye el tiempo de uso del procesador por parte del *kernel* para iniciar la aplicación, que es el costo por leerla y traerla a la memoria.

En un procesador, la diferencia entre el tiempo real y el tiempo total de procesador es:

$$\text{real} - (\text{user} + \text{sys})$$

Es la suma de todos los factores que pueden postergar la aplicación, más los costos propios del programa no atribuidos. En un ambiente de multiprocesamiento simétrico, una aproximación sería como sigue:

$$\text{real} * \text{número_de_procesadores} - (\text{user} + \text{sys})$$

Los factores que pueden disminuir el tiempo del uso del procesador son aproximadamente en este orden, por lo que se requiere analizar estos factores también.

- Tiempo de lectura/escritura al disco requerido para traer los datos y textos del programa
- Tiempo de lectura/escritura requerido para adquirir la memoria real para el uso del programa
- Tiempo del procesador utilizado por otros programas
- Tiempo del procesador utilizado por el sistema operativo

3.5.2.6 Identificación del uso intensivo del procesador por las aplicaciones

Es importante identificar los programas que están ejecutándose en el sistema operativo e identificar los recursos que utilizan. Se debe considerar que el afinamiento de las aplicaciones o programas es un tema amplio que debe de ser tomado como tema de otro trabajo de graduación, pero en esta sesión se hará énfasis en la manera de cómo detectar el uso intensivo del procesador por las aplicaciones. Por lo que se deben de ver datos como los siguientes:

- El tiempo de uso del procesador recientemente utilizado por los procesos (en unidades de ciclo de reloj).
- El total de tiempo del procesador utilizado por el proceso desde su inicio (en minutos y segundos)
- El total de tiempo del procesador utilizado por el proceso desde su inicio, dividido el tiempo transcurrido desde el inicio del proceso. Esta medida es la dependencia del uso del procesador por la aplicación. En otras palabras (tiempo del procesador por el proceso / duración del proceso) * 100, que es el porcentaje del procesador (%procesador) que ha sido asignado al proceso desde que este inicio.

Cuando se han identificado los programas de mayor uso, es necesario verificar su codificación (cuando son programas desarrollados internamente o se tiene el código fuente). Un programa típico es una mezcla variable de código de aplicación, subrutinas de librerías, y servicios del *Kernel*. Un programa que no ha sido afinado, generalmente consume más ciclos de procesador para ejecutar unas cuantas sentencias o subrutinas.

3.5.3 Soluciones a los problemas del procesador

Una manera de solucionar los embotellamientos del procesador consiste en actualizarlo por otro más rápido, si la placa base del sistema lo permite. Si se dispone de un sistema multiusuario que ejecuta aplicaciones con varios thread, puede aumentarse la potencia de proceso añadiendo más procesadores. (Si un proceso tiene varios *thread*, la potencia se aumenta añadiendo procesadores, pero si tiene sólo uno se aumenta elevando la velocidad del procesador).

No todos los procesos requieren de la misma cantidad de recursos del sistema, muchos procesos, tales como aplicaciones de bases de datos trabajan con archivos grandes, tienden a estar realizando mucha actividad con los discos, es decir requieren una configuración ideal para escribir y leer al disco.

Otros programas como compiladores requieren un uso intensivo del procesador para realizar una gran cantidad de instrucciones. Otras tareas requieren mucha memoria para leer gran cantidad de datos y manipularlos en la memoria, por lo que es de manera básica, identificar el objetivo y actividades de las aplicaciones y procesos, para poder orientar el monitoreo y realizar los cambios necesarios, según las necesidades de cada proceso.

3.5.3.1 Prioridad de ejecución

El *kernel* del sistema operativo maneja un valor de prioridad (también llamado *scheduling priority*), el cual es un valor numérico, que indica la importancia del *thread*, entre más pequeño sea este valor, más importante es el *thread* y por consecuencia tiene un mayor privilegio a que se le asigne el procesador.

Para mejorar el rendimiento de algunos procesos, o de algunas aplicaciones específicas, unas de las opciones es redefinir la prioridad del proceso, asignándole un valor que corresponda a un proceso de mayor importancia.

Motivos para manejar la prioridad de los procesos

- Compartir el procesador entre todos los *threads*
- Prevenir la inanición de cualquier *thread*
- Penalizar los *threads*
- Incrementa la discriminación continua entre *threads* sobre el tiempo.

3.5.3.2 Tiempo de asignación del procesador

El tiempo de asignación del procesador es un tiempo que tiene cada *threads* en ejecución antes de que el procesador sea asignado a otro *thread* de la misma prioridad. El control del procesador de cada *threads*, puede cambiar antes de que finalice su tiempo de asignación, debido a otros factores como una llamada de interrupción, o cuando se necesita de otro recurso del sistema y éste no está disponible inmediatamente, provocando que el *thread* asignado al procesador pase nuevamente a la cola de procesos en espera.

El tiempo de asignación del procesador es un parámetro que puede ser configurado en cualquier momento, el cual puede ser utilizado para asignar un mayor tiempo de asignación del procesador a los procesos o reducir este tiempo.

3.5.3.3 Ajuste de la carga del sistema

Distribuir los programas de forma más eficiente entre los servidores, o hacer que las aplicaciones se ejecuten a horas en las que el sistema no esté tan sobrecargado, es otra forma de mejorar el uso del procesador.

Muchas aplicaciones pueden, debido a su función, ser ejecutadas en horarios donde la carga de trabajo en general es menor. Este tipo de implementación ayuda de manera que las aplicaciones críticas tengan un rendimiento aceptable.

La distribución de la carga de trabajo, entre los diferentes servidores, es una tarea que también es importante al momento de estructurar la división de trabajo, reduce la contención de las áreas críticas y maximiza el rendimiento.

3.5.3.4 Memoria caché

La memoria caché, es una porción de memoria que funciona de manera intermedia entre el procesador y el total de la memoria real, de manera que permite ejecutar instrucciones y leer y escribir datos de manera rápida. Se denomina también CPU *caché*. Un banco de este tipo de memoria, proporciona un mayor desempeño al procesador, por tener un tiempo de acceso menor a la memoria RAM, en la cual puede ser entre cinco y seis veces menor el tiempo de acceso.

Tabla V. Parámetros para la evaluación y rendimiento del procesador

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIA
Factor de balanceo del procesador	Indicador de la actividad del procesador en tareas propias del usuario y del sistema operativo	50-55 % usuario 30-35 % sistema 10-15 % disco	Rangos más comunes, cuando la carga y el sistema responden adecuadamente
Tiempo por cambio de contexto	Tiempo estimado en el cual el procesador guarda toda la información de un proceso, y carga uno nuevo	Promedio de 10 milisegundos	Un valor mayor es un indicador de problemas del disco. Los cambios de contexto deben de ser rápidos y no mayores a éste parámetro, a no ser que el disco presenta problemas de acceso
Interrupciones por segundo	Indicador de cuantas interrupciones son atendidas por el procesador	Mínimo valor posible.	Cambios dramáticos en este indicador, son consecuencia de un mal desempeño del <i>hardware</i> , y es necesario verificar la cantidad máxima de interrupciones, con las especificaciones del fabricante

Continuación 2/3

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIA
Porcentaje de uso del procesador	Indicador que proporciona un índice de que tan ocupado está el procesador	60 % carga normal. 90 % carga alta.	Cuando este factor está por arriba del 90%, es un indicador que probablemente el procesador esté muy atareado. Cuando la carga es alta y para mejorar el rendimiento, es necesario adicionar otro ó actualizar el procesador
Longitud de colas de procesos	Longitud de procesos en espera a que el procesador pueda atenderlos	5 es el máximo valor que debería tener Un valor aceptable es de 2 unidades 0 un valor ideal	Entre menor es este número, indica que el procesador atiende los procesos oportunamente. Cuando incrementa es indicador de que el procesador hace esperar a los procesos antes de atenderlos

Continuación 3/3

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIA
Longitud de cola de procesos en espera (<i>sleep</i>).	Es otro indicador de la cola de procesos, en la cual muestra la cantidad de procesos en espera, porque un recurso diferente al procesador no está disponible	Mínimo valor posible	Cuando un proceso tiene mucho tiempo de espera por un dispositivo, o recurso diferente al procesador, se debe de verificar si el dispositivo no está disponible. Generalmente este problema es del disco
Factor de utilización del procesador por aplicación	Indica que tanto uso del procesador es requerido por las aplicaciones	Distribución uniforme	Cuando una aplicación consume más tiempo, es de determinar si es una aplicación con cierto grado de afinamiento. También es ideal examinar las políticas de prioridades de asignación del procesador.

3.6 Evaluación del rendimiento y afinamiento del disco

Los dispositivos de almacenamiento como disco, tienen una estructura jerárquica a nivel del sistema operativo, que va desde el dispositivo físico, hasta los archivos lógicos almacenados, atravesando una división lógica del disco, bloques, pistas, etc., la cual permite administrar y monitorear este tipo de recursos.

Los contadores de disco lógico miden el rendimiento de los elementos de alto nivel, como conjuntos de volúmenes. Estos contadores resultan útiles para determinar qué partición provoca la actividad del disco y ayudar a identificar la aplicación o servicio que genera las peticiones. Los contadores de disco físico muestran información acerca de discos individuales, sin tener en cuenta el modo en que son utilizados. Los contadores de disco lógico miden la actividad en las particiones lógicas de los discos, mientras que los contadores de disco físico miden la actividad en todo el disco físico.

3.6.1 Problemas con los dispositivos de almacenamiento

Los dispositivo de almacenamiento son recursos del sistema que permiten almacenar programas, archivos de textos, datos con determinado formato, datos utilizados por los programas, que generalmente no pueden permanecer en memoria, que sería ideal. Pero existen varios factores que deben de analizarse pues afectan el rendimiento para recuperar la información.

3.6.1.1 Fragmentación

Lógicamente se puede pensar en un archivo en el sistema operativo como una cadena de caracteres secuenciales, aunque físicamente esté dividido en unidades lógicas distintas, bloques de datos no contiguos, etc., que en resumen en un momento determinado puede estar fragmentado y no solamente un archivo, si no todo los archivos almacenados en el dispositivo físico. Un dispositivo de almacenamiento físico, con una alta fragmentación, muy probablemente tenga problemas para acceder a la información de manera secuencial, la búsqueda requerirá de un mayor tiempo para la respuesta.

Con la información almacenada en el disco, se debe determinar la fragmentación existente en los archivos, debido a que a mayor fragmentación, mayor es el tiempo para recuperarla.

3.6.1.2 Lectura de archivos

Cuando se lee un archivo, este es cargado a la memoria por el VMM (*Virtual Memory Manager*) a bloques de memoria y comúnmente se cargan unos cuantos bloques del archivo, principalmente cuando es demasiado grande. La información almacenada en la memoria es mucho más fácil de acceder, y cuando los datos son modificados, éstos son actualizados en el disco, en los bloques correspondientes.

El VMM trata de anticipar los requerimientos de páginas de un archivo secuencial, con base en los patrones de acceso del archivo por las aplicaciones. Cuando una aplicación solicita información de páginas continuas de un determinado archivo, el VMM asume que la aplicación continuará solicitando información del archivo de manera secuencial, por lo que programa leer bloques de manera anticipada para tenerlas en memoria. Estas lecturas son intercaladas con el procesamiento del programa, y tendrán los datos disponibles tan pronto como el programa los necesite. El número de páginas a ser leídas de manera anticipada está determinado por dos parámetros que indican la cantidad de páginas mínimas y máximas a leer de manera anticipada.

Cuando el programa lee bloques fuera de la secuencia, entonces el VMM deja de leer bloques secuenciales y reinicia nuevamente la lectura de bloques secuenciales, únicamente si detecta que el programa lee nuevamente el archivo de manera secuencial.

3.6.1.3 Escritura de archivos

Para incrementar el rendimiento de escritura, se puede limitar la cantidad de páginas de archivos en memoria, de esta manera se reduce también la fragmentación del archivo y la sobrecarga del sistema por acceso a lectura/escritura del disco. El sistema de archivos divide los archivos en particiones de un tamaño fijo de pueden variar dependiendo del sistema operativo. Las páginas de la memoria no son escritas al disco, hasta que el programa escriba los primeros bytes de cada partición.

Estas páginas de memoria de archivos permanecen en la memoria hasta que se necesiten reutilizar los bloques de la memoria. Si son requeridos nuevamente, mientras permanecen en la memoria no es requerido el acceso al disco para la lectura.

Si una cantidad considerable de bloques de memoria de archivos permanecen en memoria, automáticamente se inicia un proceso para escribirlos en el disco. Al realizarse este proceso ocurre un utilización anormal del disco que genera problemas de rendimiento, como consecuencia de esto, para minimizar la utilización anormal del disco, los sistemas definen dos parámetros para indicar la cantidad máxima y mínima de bloques de memoria para archivo que pueden permanecer (tal y como lo expusimos en el uso de la memoria, de las sesiones anteriores) antes que el proceso automático del sistema operativo escriba los bloque al disco, con estos parámetros se distribuye la carga de los bloques lo que además permite distribuir la actividad de lectura/escritura del disco.

3.6.1.4 Contención

La contención de disco ocurre cuando múltiples procesos tratan de acceder el mismo disco simultáneamente. La mayoría de discos tienen limitaciones en el número de accesos concurrentes y la cantidad de datos que pueden transferir por segundo. Cuando dos procesos tratan de acceder al mismo tiempo, generalmente un proceso tiene que esperar para poder acceder al disco.

3.6.1.5 Atributos del sistema de archivo que afectan el rendimiento

Cuando más tiempo tiene el sistema de archivos en uso, mucha más fragmentación existirá, con la asignación dinámica de recursos, los bloques empiezan a estar más y más dispersos, lógicamente los archivos contiguos son fragmentados, y los volúmenes lógicos contiguos también son fragmentados.

Cuando los archivos son utilizados se producen los siguientes efectos

- El acceso secuencial ya no es secuencial.
- El acceso aleatorio es mucho más lento
- El tiempo de acceso es dominado por la búsqueda de mayor tiempo.

Sin embargo, cuando el archivo está en memoria, estos efectos disminuyen. El rendimiento del sistema de archivos es también afectado por las consideraciones físicas, tales como.

- Tipo de disco y número de adaptadores
- Cantidad de memoria a ser utilizada para caché de archivos
- Cantidad de acceso a archivos remotos y archivos locales
- Patrones y cantidad de acceso de archivo por aplicación

3.6.1.5.1 Definición de los bloques de datos en el disco

Otro efecto adverso en la actividad de disco es el número de operaciones de lectura/escritura. Para un archivo con un tamaño de 4Kb es guardado en un bloque de 4Kbytes, y únicamente se requiere de una operación de escritura y otra operación para lectura del archivo. Si el tamaño de los bloques de almacenamiento en el disco fueran de 512 bytes, se necesitarían ocho bloques para asignar este archivo, y para poder leer y escribir el archivo, se necesitarían varias operaciones de lectura/escritura (buscar en el disco, transferir datos y asignar la actividad). Por consiguiente, para un sistema de archivos que utilizan fragmentos de bloque de tamaño de 4Kb, el número de operaciones de lectura/escritura puede ser mucho menos que para el sistema en el cual se utiliza en bloque de tamaño menor.

La consideración de crear fragmentos de bloques menores o de un determinado tamaño para el sistema de archivos, dependerá de la política de fragmentación que se desea del disco, teniendo en cuenta que cuando se requiere de un bloque adicional y no se utilice de manera completa, el espacio será desperdiciado.

3.6.1.5.2 Compresión de datos

Cuando el sistema de archivos utiliza propiedades y atributos de compresión de datos, debe de tomarse en cuenta que todos los datos automáticamente serán comprimidos, lo cual representa algunos inconvenientes para el rendimiento del sistema de archivos.

- Los datos utilizan un mayor tiempo del procesador para comprimir y descomprimir, lo cual se refleja en una degradación del sistema de archivos.
- Para llevar a cabo la compresión y descompresión de datos o información, se requiere de la utilización del procesador, el cual también es un recurso muy importante.
- Los datos se requieren lo más pronto posible, principalmente cuando son de un sistema de tiempo real, lo cual no hace que la compresión sea factible por los tiempos de respuesta.
- Cuando se utilizan los datos de un sistema comprimido, se necesita espacio adicional para este proceso el cual se traduce en fragmentación del disco.

3.6.2 Parámetros del disco a evaluar

Para la evaluación y monitoreo de los dispositivos de almacenamiento, tales como el disco que afectan de manera significativa el rendimiento de las aplicaciones y del sistema operativo, existe un gran conjunto de parámetros que pueden ser evaluados y que en conjunto con otros valores de diferentes áreas, como utilización de procesador, memoria, etc., proporcionan información que es de utilidad para determinar las problemas que pueden ser mejorados y la manera que puede tener un mejor desempeño en el área de los medios de almacenamiento como los discos.

Algunos de los principales parámetros que se detallan a continuación, pueden describir operaciones y valores que permiten conocer e identificar áreas del rendimiento de los discos, tanto desde la parte física, hasta la parte lógica de éstos.

3.6.2.1 Tiempo de utilización del disco

Generalmente el disco está ocupado o bien está disponible para su utilización en alguna actividad, ya sea de lectura o escritura por algún proceso. Debe de comprobarse el porcentaje de utilización del disco y el porcentaje de disco inactivo, en total estos dos contadores no deben de ser mayor al 100%. Si este valor se encuentra constantemente en las cercanías del 100%, el sistema utiliza intensamente el disco. Si el disco se encuentra constantemente ocupado y se ha formado una larga cola, es posible que se haya producido un embotellamiento en el disco. En condiciones normales, este valor no debería sobrepasar el 50% de utilización.

Pero cuando el porcentaje de utilización del disco es de manera continua mayor a 85%, se tiene un embotellamiento por disco.

$\% \text{utilización de disco} > 85 \Rightarrow \text{embotellamiento.}$

3.6.2.2 Tiempo de espera por disco

Un alto porcentaje de tiempo de espera `%iowait` (*input output wait*, que es un indicador del tiempo del procesador sin uso mientras espera por el disco) indica que el sistema tiene una aplicación con problemas de acceso al disco, poca capacidad de memoria, o una configuración del subsistema de entrada/salida ineficiente. Por ejemplo, la aplicación con problemas puede realizar una gran cantidad de operaciones de entrada/salida a disco, pero no realiza mayor operación y cálculo con los datos.

Si el `%iowait > 30%` en un sistema de desarrollo, en el cual son pocos procesos y programas los que se ejecutan, se debe de chequear el punto en donde este indicador tiene la irregularidad, observando los parámetros `%wfs` (*waiting for file system*) y `%swp` (*waiting for swap*), con los cuales se puede precisar exactamente, cual es el motivo de espera del sistema.

En un sistema NFS que está haciendo de servidor de archivos (*file server*) y está atendiendo peticiones de clientes, si el cliente tiene `%iowait > 80` y `%wfs > 90`, entonces el sistema presenta un problema de disco.

3.6.2.3 Transferencia de datos por segundo

Una transferencia es una requisición de lectura/escritura a nivel del manejador del dispositivo hacia el disco físico. Múltiples requisiciones lógicas pueden ser combinadas dentro de una requisición de lectura/escritura hacia el disco. Una transferencia es de un tamaño indeterminado. La transferencia de datos es en sí la cantidad de datos leídos y escritos en Kbytes por segundo.

La suma de los datos leídos más los datos escritos al disco, dividido por el intervalo de tiempo medido, proporciona la cantidad de la transferencia de datos.

Cuando en una evaluación no se cuenta con muchos datos estadísticos, se puede utilizar la relación entre utilización de disco y tasa de transferencia de datos, para encontrar patrones y relaciones de datos en la utilización del disco.

3.6.2.4 Media en segundos por transferencia

Este parámetro muestra la cantidad de tiempo que le lleva a un disco completar una petición. Un valor alto podría indicar que el controlador de disco se encuentra continuamente intentando acceder al disco como consecuencia de los errores. Para la mayoría de los sistemas, un valor de 0.3 segundos o superior indica una media de tiempo de transferencia de disco alta.

3.6.2.5 Cola de actividades del disco

La media de la cola de actividades de un disco, es un buen indicativo de embotellamiento por utilización, este es un parámetro que no debería de exceder de 3 unidades, porque de ser así y de manera continua, se tiene problemas por rendimiento. Durante períodos de intenso acceso al disco, resulta habitual la existencia de solicitudes en cola, pero si aparecen constantemente, el disco no es lo suficientemente veloz para el tipo de trabajo que se realiza, por lo que de ser posible debería de cambiarse por otro más rápido o de diferente arquitectura de almacenamiento.

Cuando se disponen de más de un disco, el parámetro general para todas las instancias, no debería de ser más a

2 * el número de ejes que constituyen el disco físico

La excepción son los dispositivos de disposición redundante de discos independientes (*RAID*), que normalmente tienen más de un eje. Es necesario observar este valor durante varios intervalos, dado que se trata de un valor instantáneo.

3.6.2.6 Tiempo por espera de lectura/escritura del disco

En un sistema de aplicaciones con un gran porcentaje de espera en lectura/escritura al disco puede estar relacionada a la carga de trabajo. En un sistema con muchos procesadores, algunos estarán ejecutando procesos mientras otros esperan por lectura/escritura en el disco. En estos casos, el porcentaje de espera por lectura/escritura (*%iowait*) puede ser pequeño o casi cero porque los procesos en ejecución disimulan un poco el tiempo de espera. A pesar de que *%iowait* es bajo, un cuello de botella puede limitar el rendimiento de las aplicaciones y del sistema en general.

Si se monitorea que no existe una situación de uso del procesador o de procesadores comprometidos, y el tiempo de espera por lectura/escritura al disco (*%iowait*) es mayor que el 20%, puede existir una actividad de lectura/escritura que compromete el rendimiento del disco. Esta situación puede ser causada por una paginación excesiva debido a la carga de la memoria física.

Una carga de disco desbalanceada, datos fragmentados o utilización de formatos en los datos afecta el rendimiento del disco.

3.6.2.7 Desempeño de utilización del disco

Es el porcentaje de tiempo que un disco físico ha estado activo (utilización de ancho de banda para el manejador del disco) o en otras palabras, el tiempo total del disco para actividades sobresalientes. Un manejador de disco está activo durante la transferencia de datos y procesamiento de comandos, tal como una búsqueda de una nueva localización. Este porcentaje es directamente proporcional a la contención del recurso e inversamente proporcional al rendimiento. Cuando el uso del disco incrementa, el rendimiento se reduce y el tiempo de respuesta se incrementa.

En general, cuando la utilización del disco excede en un 70%, entonces los procesos deben de esperar más de lo necesario para realizar sus actividades, pues por lo general los procesos son bloqueados o bien cambiados a un estado de inactivos mientras se realizan actividades del disco.

Para corregir este problema de exceso de utilización del disco, se debe de comparar los manejadores ocupados y disponibles y hacer una transferencia de carga a un manejador que esté disponible.

3.6.2.8 Fragmentación del disco

Cuándo existe una alta actividad del disco con una alta transferencia de datos, no es motivo de preocupación de rendimiento. Pero cuando existe una alta utilización del disco y una baja transferencia de datos, es posible que exista una fragmentación del volumen lógico, *file system* o en algún archivo individual.

3.6.2.9 Eficiencia de espacio y secuencialidad en el disco

Una alta eficiencia en espacio significa que los archivos son menos fragmentados y probablemente proveen un mejor acceso secuencial. Una alta secuencialidad indica que los archivos son contiguos unos a otros y esto probablemente haga el acceso secuencia de archivos mucho mejor.

Eficiencia de espacio = $\frac{\text{total de número de fragmentos utilizados para almacenar archivos}}{(\text{máximo fragmento de dirección física} - \text{mínimo fragmento de dirección física} + 1)}$

Secuencialidad = $\frac{(\text{total de número de fragmentos} - \text{numero de fragmentos agrupados} + 1)}{\text{total de números de fragmentos}}$

3.6.3 Soluciones para el afinamiento de los discos

El afinamiento del sistema de archivos, es uno de los componentes que afecta drásticamente el rendimiento de los sistemas operativos, muchas veces afinar estos dispositivos puede ser muy tedioso y consume demasiado tiempo, por consiguiente se debe de analizar la cantidad de dispositivos, tipos de dispositivos, tamaños, unidades lógicas, espacio de los bloques, previo a su instalación.

En el afinamiento de los medios de almacenamiento, se debe de determinar los archivos más utilizados, *file system* y volúmenes lógicos, luego preguntar si los *file system* o unidades lógicas pueden ser ubicado en otra unidad física, o mejor aún en varias unidades.

Determinar cuales son los archivos más utilizados y la ubicación, ya sean locales o remotos.

Con la paginación del disco determinar si es posible que dominen el uso por este problema.

Con cuanto espacio de memoria se disponemos para tener páginas de disco en caché, o que cantidad de acceso de archivos a disco utilizan los procesos de manera asincrónica (sin caché).

Identificar los volúmenes físicos con una alta utilización y concluir si la unidad o el adaptador de lectura/escritura es el causante de un embotellamiento en el disco.

3.6.3.1 Embotellamiento por lectura o escritura al disco

Para el problema de embotellamiento de lectura/escritura, la solución consiste en mejorar la eficiencia del subsistema. Esta solución puede ser de diferentes formas, algunas soluciones típicas incluyen

- Limitar el número de volúmenes lógicos activos y sistema de archivos almacenados en un disco físico en particular. La idea es balancear la carga de lectura/escritura de archivos entre todos los manejadores del disco físico.
- Distribuir los volúmenes lógicos entre múltiples discos físicos. Esto es particularmente útil cuando un determinado número de archivos son leídos.
- Crear múltiples áreas para cálculos temporales y manejo de archivos temporales, preferiblemente en dispositivos de escritura rápida.
- Cuando se determina que existe una alta actividad de carga de trabajo de lectura/escritura y no es distribuida uniformemente en los manejadores del sistema de archivo, y la utilización de uno o mas manejadores de discos está entre 70-80 por ciento o más, debe de considerarse reorganizar el sistema (*file systems*), tal como crear un *back up* y restaurarlo para reducir la fragmentación. La fragmentación causa que el manejador realice una excesiva búsqueda y puede ser una gran sobrecarga del tiempo de respuesta.

- Cuando una pequeña cantidad de archivos son leídos una y otra vez, se debe considerar agregar memoria física que permitiría a estos archivos estar en la memoria mucho más tiempo y ser más eficiente.
- Cuando el acceso de la carga de trabajo tiene características de lectura de patrones de información y son predominantemente secuenciales e involucran a múltiples manejadores del disco, debería considerarse agregar uno o más adaptadores de disco. Además, puede ser más apropiado considerar la creación de un volumen lógico de arreglo de discos.
- Cuando la carga es predominantemente aleatoria, debe de considerarse agregar discos y distribuir el acceso a archivos aleatoriamente a través de más manejadores.
 - Utilización de dispositivos de caché de escritura rápida.
 - Utilización de escritura sincrónica.
- Añadir un controlador *RAID* más rápido, añadir más discos a una configuración *RAID* (distribuir los datos entre varios discos físicos mejora el rendimiento, sobre todo en las lecturas)
- Añadir más memoria para aumentar el tamaño del caché.

3.6.3.2 Archivo de paginación

El archivo de paginación debe de tener el tamaño correcto (tal como se definió en las soluciones de memoria, que equivale al tamaño total de memoria más el tamaño del encabezado del archivo), así como de crear múltiples archivos de paginación para reducir la paginación excesiva. También puede dividirse el archivo de paginación entre varios discos de velocidades parecidas para incrementar el tiempo de acceso.

De ser posible puede adicionarse más memoria para reducir la cantidad de lecturas y escrituras de disco, esto permitirá reducir el paginación e incluso el área de *swap*.

3.6.3.3 Utilización de adaptadores del disco

La lucha por la utilización del disco, volúmenes lógicos y rendimiento del sistema de archivos, muchas veces indica que uno o más manejadores de disco en el sistema, pueden mejorar el rendimiento de lectura/escritura. Esto no siempre es cierto porque existe un límite de la cantidad de datos que pueden ser manejados por el adaptador del disco. El adaptador del disco puede ser en un determinado momento el cuello de botella.

Si todos los manejadores del disco están en un único adaptador y el sistema de archivos está en un volumen físico separado, se puede beneficiar el uso de múltiples adaptadores de disco. El rendimiento mejora dependiendo del tipo de acceso a la información.

3.6.3.4 Striping

El concepto de *striping*, consiste en dividir un archivo grande en porciones pequeñas y almacenar en diferentes discos cada una de estas partes pequeñas. Esto permite que múltiples procesos concurrentemente puedan acceder al mismo tiempo a un mismo archivo en diferentes discos.

Striping es ideal cuando los accesos a los archivos son de manera aleatoria. El uso de esta técnica le puede permitir tener hasta un 50% de mejora en el rendimiento de lectura/escritura.

Striping trabaja mejor en discos que están en diferentes controladores. Los discos *striping* proporcionan más espacio sin disminuir el rendimiento potencial, pero corre el riesgo de que si se pierde uno de los discos con *striping*, se pierde toda la información del sistema de archivos, puesto que los datos están intercalados a través de todos los discos

3.6.3.5 Sugerencias para el ajuste y progresión de la actividad del disco

Otras sugerencias para mejorar el rendimiento del disco son

- Instalar el último controlador *software* para los adaptadores de red para mejorar la eficiencia de acceso al disco.

- Instalar discos adicionales o actualizar el disco duro a uno más rápido. Hay que actualizar también al mismo tiempo el controlador del disco y el bus de datos.
- En servidores, crear volúmenes repartidos en varios discos físicos para incrementar la productividad.
- Distribuir las aplicaciones entre varios servidores para equilibrar la carga de trabajo.
- Optimizar el espacio del disco ejecutando el desfragmentador.
- Aislar las tareas que requieran un uso intensivo de entrada/salida para separar los controladores de disco o discos físicos y poder equilibrar la carga de trabajo del servidor.

3.6.3.6 Agregar más discos para mejorar el rendimiento

Si no mejora el rendimiento del disco aún después de realizar el afinamiento y mejoras sugeridas, debe de considerarse agregar más discos, controladores e incluso memoria.

Para determinar que tanto o tipo de *hardware* puede adicionarse es necesario consultar siempre las especificaciones del *hardware* del equipo al que se desea agregar más recursos, y de esta manera seleccionar el disco o controlador correcto para las necesidades particulares que se tengan.

Tabla VI. Parámetros para la evaluación y rendimiento del disco

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIAS
Tiempo de utilización del disco	Indicador del porcentaje de actividad del disco	50 % utilización normal. 85% utilización alta del disco	El tiempo de utilización es un factor que determina que tanta actividad tiene el sistema operativo a nivel del disco
Tiempo de inactividad del disco	Indicador del porcentaje de inactividad del disco	50% inactividad normal	El tiempo de inactividad más el tiempo de utilización del disco debe de sumar 100%
Porcentaje de espera por el disco	Indicador del tiempo del procesador sin uso mientras espera por el disco	Máximo 30%	Cuando tenemos más de este porcentaje, hay que examinar las áreas de <i>swap</i> y de paginación, para determinar el bajo rendimiento del disco

Continuación 2/5

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIAS
Transferencia de datos por segundo	Cantidad de datos leídos y/o escritos del disco, entre el intervalo de tiempo transcurrido para realizar dicha actividad	Mayor posible	Cuando mayor sea la cantidad de datos que pueda leer y escribir, mayor es la eficiencia de los medios de almacenamiento
Media en segundo por transferencia	Tiempo promedio que necesita el disco para completar una petición del sistema operativo	Menor posible.	Cuando menor sea el tiempo para completar una petición, más eficiente es el sistema de discos.
Cola de actividad del disco	Promedio de procesos en espera por atención del disco	Máximo de 2 unidades por disco físico	Es un indicador de eficiencia del disco. Si este parámetro es mayor a este valor, es recomendable cambiarlo por un disco más reciente o de mayor velocidad

Continuación 3/5

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIAS
Tiempo de espera de lectura / escritura del disco	Porcentaje del tiempo promedio que los procesos esperan para poder realizar una operación de lectura/escritura al disco	Máximo 20%	Cuando el porcentaje de tiempo de espera es mayor a este parámetro, el acceso al disco puede ser un factor de bajo rendimiento en el sistema operativo
Desempeño de utilización del disco	Es un porcentaje de la utilización de disco	Máximo 70%.	Generalmente se espera que los recursos estén siendo bien utilizados, pero en el caso de los discos, cuando están siendo muy utilizados, es probable que algunos procesos tengan que esperar, por lo que es recomendable que se agreguen más dispositivos para distribuir la carga y el factor de utilización disminuya

Continuación 4/5

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIAS
Fragmentación	Este es un indicador que proporciona un porcentaje de la fragmentación de los dispositivos de almacenamiento. Es decir que tan esparcida está la información	Máximo 20%	Cuando se tiene una aplicación que utiliza información de manera secuencial de los discos, es recomendable que el porcentaje de fragmentación sea mínimo
Tamaño del archivo para paginación	Total de todos los archivos utilizados para paginación	Memoria física más 12 Kbytes por cada archivo utilizado para paginación.	Tamaño de memoria física más tamaño de encabezado de disco (12 Kbytes) en promedio. Con excesiva paginación es necesario adicionar más memoria o bien, utilizar discos más rápidos

Continuación 5/5

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIAS
Utilización del archivo de paginación.	Porcentaje de utilización de área de disco reservada para paginación.	Menor posible	Cuando la utilización del archivo de paginación es demasiado alto, es probable que sea necesario adicionar memoria física.
Porcentaje de utilización de paginación.	Porcentaje de utilización de área definida para paginación.	Menor posible	Es recomendable que el porcentaje de uso de paginación sea el mínimo posible, esto implica tener suficiente área de memoria física.

3.7 Evaluación del rendimiento y afinamiento de la red

Cuando aumentan los problemas de rendimiento, puede ser que el sistema no sea el culpable, y en realidad el culpable sobre el bajo funcionamiento se encuentre fuera, es decir el bajo rendimiento es como consecuencia de otro sistema con el cual se comparten recursos por medio de una red. Una manera fácil de identificar si la red está afectando todo el rendimiento es comparar aquellas operaciones que competen a la red con aquellas que no involucran trabajo en red. Cuando se ejecuta un programa que realiza una considerable cantidad de lecturas y escrituras y está lento en su desempeño, pero todos los demás procesos se aprecia que están ejecutándose de manera usual y son por decirlo así locales, entonces es probablemente un problema de red.

3.7.1 Problema de la red

Para determinar dónde residen los embotellamientos de red y cómo eliminarlos, es necesario entender el tipo de carga que generan los clientes, los componentes clave de la arquitectura de red que se ven involucrados y el tipo de red física y el protocolo de red (*Ethernet*, *NetBEUI*, etc.) que se utiliza.

Algunos problemas potenciales de embotellamiento de red pueden ser causados por las siguientes características

- Interfase de red del cliente
- Ancho de banda de red
- Topología de red
- Interfase de red del servidor
- Carga del procesador del servidor

- Uso de memoria del servidor
- Ancho de banda del servidor
- Configuración ineficiente

Las herramientas de monitoreo pueden medir estadísticas de red y proporcionar una variedad de información, pero únicamente una parte de esta información está relacionada al afinamiento del rendimiento, la cual se identifica en esta sesión y los parámetros necesarios para mejorar el desempeño y rendimiento de las redes.

3.7.2 Parámetros para medir y evaluar el rendimiento de las redes

A continuación, se mencionan algunos parámetros que son de utilidad para evaluar el rendimiento de la red, y posteriormente se determinarán unos lineamientos para saber el comportamiento de la red.

- Cantidad máxima de unidades de transmisión de los paquetes en bytes utilizando la interfase.
- Número total de paquetes recibidos
- Número total de errores ingresados. Por ejemplo, paquetes mal formados, errores de chequeo de sumas de tamaño de paquetes, o insuficiente espacio del *buffer* en el manejador de dispositivo de red.
- Número total de paquetes transmitidos
- Número total de errores de salida. Por ejemplo, una falla en la conexión del adaptador del *host* local.
- Número de colisiones de paquetes detectados.

Estos parámetros proporcionan información que permite identificar embotellamientos y además darán una idea de qué está causando un embotellamiento a nivel de la red de manera que se pueda corregir aspectos de rendimiento.

3.7.2.1 Paquetes recibidos

Si el número de errores durante el ingreso de paquetes es mayor al 1% del total de paquetes ingresados.

$$Ierrs > 0.01 \times Ipkts.$$

Entonces verifique la disponibilidad de memoria

Si el número de errores durante el envío de paquetes es mayor al 1% del total del número de paquetes enviados.

$$Oerrs > 0.01 \times Opkts.$$

Entonces incremente el tamaño de la cola de envío de la interfase.

Si la tasa de colisiones es mayor que el 10%

$$Coll / Opkts > 0.1$$

Entonces existe una alta utilización de la red, y una reorganización o partición puede ser necesaria.

3.7.2.2 Colas de recepción y envío de paquetes

Si el contador de la longitud de cola de salida, que mide la longitud de la cola de paquetes de salida de un adaptador, tiene un valor comprendido entre 1 y 2 es un valor aceptable.

3.7.2.3 Colisiones

Si el número total de colisiones es mayor del 10% del total de paquetes transmitidos, entonces se debe investigar un poco más, pues probablemente exista una sobrecarga en la red.

Si el número de colisiones / número de paquetes transmitidos > 0.1

Entonces es un indicador de sobrecarga de la red.

3.7.2.4 Fragmentación de paquetes

La fragmentación de paquetes implica una sobrecarga adicional en el procesador por lo que es importante determinar la causa. Algunas aplicaciones, por su propia naturaleza, pueden estar enviando una gran cantidad de paquetes fragmentados, cuando esto ocurre, puede ser que el MTU¹⁰ (unidad máxima de transmisión) utilizado sea de un tamaño diferente al del sistema.

¹⁰ Unidad máxima de transmisión, es el tamaño máximo de paquete, en bytes que puede manejar una interfase en particular.

3.7.2.5 Redes *Ethernet*

Para chequear si una red *Ethernet* está sobrecargada, es necesario calcular

$$\left(\begin{array}{l} \text{número máximo errores por colisiones} + \\ \text{número de transmisiones fallidas} \end{array} \right) / \text{cantidad de paquetes transmitidos.}$$

Si el resultado es mayor de 5%, es posible que se necesite reorganizar la red para balancear la carga.

3.7.2.6 Redes con protocolo IP (*Internet Protocol*)

El protocolo IP es un formato de conexiones de ordenadores con la red de Internet. Protocolo a nivel de capa de red, brinda funciones de direccionamiento en Internet, especificación del tipo de servicio, fragmentación y reensamble.

Parámetros de interés a monitorear en este tipo de protocolo son

- Número total de datagramas IP recibidos.
- Error de chequeo de paquetes o fragmentos eliminados. Esta opción verifica los paquetes recibidos por medio de un número agregado a cada paquete enviado, el cual al no concordar asume que el paquete está corrupto o fragmentado, y esto es debido a que la red está fragmentando los paquetes o bien la cola del manejador del dispositivo que recibe los paquetes es demasiado pequeño.

- Número total de fragmentos recibidos
- Cantidad de paquetes eliminados luego de que el contador de vigencia ha expirado. Debido al tráfico de red que impide que los paquetes lleguen antes a su destino.
- Número de datagramas IP que han sido creados y enviados por el sistema.
- Número de fragmentos creados, cuando los datagramas IP son enviados.

Observar estas estadísticas, y ver el promedio de paquetes recibidos y fragmentos recibidos como lineamiento de MTU. Si el 10% de los paquetes o más están siendo fragmentados, es importante determinar el motivo de esta cantidad de fragmentación de paquetes. Una gran cantidad de fragmentación indica que el protocolo a nivel de la capa IP en el servidor remoto, está pasando datos a la IP en un tamaño mucho mayor a la cantidad MTU para la interfase.

Los *gateway/routers* en la red pueden tener el MTU de un tamaño mucho menor que otros nodos. La misma lógica puede aplicar a los paquetes enviados y fragmentados.

3.7.2.7 Protocolo UDP (*User Datagram Protocol*)

Protocolo de datagramas de usuario, que no necesita una conexión directa entre el destinatario y el remitente, y permite el envío de datos a través de la red. Los datagramas son enviados sin causa o envío de recibo o garantía de entrega y requiere que el procesamiento y retransmisión de errores sean manejados por otros protocolos.

Parámetros de interés a monitorear en este tipo de protocolo.

- Error de chequeo de validación de paquetes, que ocurre por problemas de hardware o fallas en el medio de transmisión.
- Número de datagramas UDP recibidos debido a que el puerto del destinatario no fue posible abrirlo. Cuando la cantidad de datagramas rechazados por este problema es alto, es necesario investigar como el manejador es administrado.
- El número de veces en que el *buffer* de los *sockets* es insuficiente, esto es debido a un desbordamiento de datos el cual es a causa de una insuficiente transmisión y recepción de los *sockets* UDP. Es necesario verificar la cantidad de procesos que administran la recepción de los *sockets*, el tamaño de *buffer* de cada *socket* y el parámetro del espacio reservado para recibir mensajes.

3.7.2.8 Protocolo TCP

Protocolo de control de transmisión de datos. Protocolo de la capa de transporte orientado a la conexión que proporciona una transmisión confiable de datos bidireccional.

Estadísticas de interés para este tipo de protocolo son:

- Paquetes enviados
- Datos empaquetados
- Datos empaquetados transmitidos nuevamente
- Paquetes recibidos
- Paquetes completamente duplicados
- Intervalos de retransmisión

Las estadísticas TCP comparan el número de paquetes enviados con el número de paquetes de datos retransmitidos. Si el número de paquetes retransmitidos es alrededor del 10 al 15 por ciento del total de paquetes enviados, entonces TCP está experimentando un alto tráfico de red, el cual provoca que los mensajes de paquetes recibidos por el destinatario no lleguen antes del tiempo de espera (consentimiento de paquetes llegados: ACKs). Un embotellamiento en el nodo de recepción o problemas generales de red pueden además causar la retransmisión de paquetes de TCP, los cuales incrementan el tráfico en la red, agregando además problemas de rendimiento.

Comparar el número de paquetes recibidos con el número de paquetes completamente duplicados. Si un nodo que envía paquetes por medio de TCP está enviando paquetes nuevamente antes de recibir de parte del nodo destinatario la señal de recepción de datos, entonces hay una retransmisión de paquetes. La duplicación ocurre cuando el nodo que recibe, eventualmente recibe todos los paquetes retransmitidos. Si el número de paquetes duplicados excede el 10 por ciento, el problema puede repercutir en una mayor cantidad de tráfico en la red o en un embotellamiento en el nodo que recibe.

3.7.3 Soluciones a los problemas de la red

Si se determina que un subsistema de la red se halla con problemas de embotellamiento, se pueden realizar diversas medidas para solucionar el problema, entre las cuales se mencionan por ejemplo, cambiar controladores por unos más recientes, sustituir los adaptadores por otros mejores o añadir más adaptadores para segmentar la red de manera que se pueda aislar el tráfico de los diversos segmentos. También debe comprobarse el rendimiento general de la red y mejorar los componentes de la capa física, como concentradores y conmutadores, para verificar si la limitación se encuentra en el tendido de la red e incluso distribuir la carga del proceso entre diversos servidores. Otras soluciones pueden ser modificar el tamaño de los protocolos, para que tengan una mayor cantidad de datos por cada paquete enviado y recibido.

La configuración óptima de los parámetros de comunicación que pueden ser mejorados por medio de un afinamiento, varía con el tipo de LAN (*Local Area Network*), como también de las características de comunicación de lectura/escritura del sistema predominante y del programa de aplicación.

Para poder identificar problemas de red, se deben realizar ciertas pruebas como enviar paquetes a sí mismo y tomar el tiempo que tardan en retornar estos paquetes por la red. Un comando común a esto es PING (*Package InterNet Groper*) buscador de paquetes en Internet.

Este tipo de prueba es ideal para lo siguiente

- Determinar el estado de la red en varios *host* foráneos.
- Rastrear y aislar *hardware* y *software* con problemas
- Verificar, medir y administrar la red.

Con este comando se puede especificar el número de paquetes enviados por la red, longitud de los paquetes, lo cual es muy útil para chequear la fragmentación y reagrupación de los mismos. Aún cuando este comando no esté propiamente como una opción de monitoreo, su información es muy útil para evaluar parámetros de rendimiento.

Cuando se trata en inundar de paquetes a la red a un destino, para medir y evaluar el rendimiento se deben tener presente ciertas consideraciones como

- El envío de paquetes agrega carga al sistema
- Monitorear el estatus de la interfase de la red durante el experimento. En el cual puede describir que el sistema elimina paquetes durante el envío.
- Monitorear otros recursos, por ejemplo, el *buffer*, la cola de paquetes enviados y recibidos. Colocar una carga de datos en el sistema de destino para sobrecargarlo puede ser difícil, pues el propio sistema es sobrecargado antes del que se desea medir.

- Se debe considerar la relatividad de los resultados. Si se desea monitorear o hacer pruebas con un determinado sistema, se debe tener en cuenta realizar experimentos en otro sistema para comparar, porque la red o el *router* (dispositivo que conecta dos redes de área local) puede tener problemas.

Al momento de llevar a cabo un afinamiento sobre la red, se puede seleccionar un objetivo primario de afinamiento, el cual puede ser entre rendimiento y uso de memoria. Algunas recomendaciones aplican a un objetivo y otros aplican a ambos.

3.7.3.1 Tamaño de los bloques de datos

Recomendación de diferentes tamaños de bloque de datos para diferentes dispositivos y adaptadores.

Tabla VII. Tamaños de bloques para dispositivos y adaptadores de la red.

Nombre del dispositivo	Aplicación del tamaño de bloque
<i>Ethernet</i>	Múltiplos de 4096
<i>Token-ring</i> (4 Mb)	Múltiplos de 4096
<i>Token-ring</i> (16 Mb)	Múltiplos de 4096
FDDI (tcp)	Múltiplos de 4096
SOCC (tcp)	28672 bytes
HIPPI	65536 bytes
ATM	Múltiplos de 4096

3.7.3.2 Maximizando el rendimiento

Utilizar las siguientes recomendaciones para maximizar el rendimiento

Cuando se trata de mejorar el tiempo de respuesta según las características del protocolo, se pueden utilizar estos consejos.

- Para maximizar el número de transacciones por segundo, utilizar el mínimo tamaño de mensajes posibles.
- Para maximizar el número de bytes por segundo, utilizar mensajes que sean al menos de 1000 bytes e igual o menor a múltiplos de 4096 bytes (4Kb).
- Si la solicitud y respuesta son de tamaño fijo y caben en un datagrama, utilizar protocolo UDP.
 - Si es posible, configurar el tamaño de las escrituras igual a lo siguiente. (en múltiplos del tamaño de MTU menos 28 bytes para permitir el estándar IP y encabezado de UDP)
 - En general, es mucho más eficiente para las aplicaciones escribir mensajes largos y que hayan sido fragmentados y reensamblados por el IP, que la aplicación tenga que escribir muchas veces.
 - En cualquier momento de ser posible, utilizar subrutinas para asociar una dirección con el *socket* UDP. Esto puede no ser posible en un servidor que está comunicándose con un número de clientes a través de un único *socket*.

- Si la solicitud o respuesta son de tamaño variable, utilice TCP con opciones de no-retardo. Las mediciones indican que los encabezados de TCP comparados con UDP son insignificantes, especialmente si la optimización para escritura es utilizada.
 - Para evitar la copia de datos en el kernel, hay que realizar escrituras de un tamaño de al menos 512 bytes.
 - Realizar escrituras de al menos del tamaño o múltiplos del parámetro MTU. Esto va a evitar el envío de segmentos (paquetes) con una cantidad mínima de bytes.

Para tener un mejor control de fluidez de los datos se puede realizar los siguientes consejos.

- TCP provee de un alto rendimiento sobre UDP y asegura de manera confiable la entrega de los paquetes.
- La escritura debe de ser en múltiplos de 16384 bytes (16Kb). Si es posible, la escritura debería de ser del tamaño del parámetro MSS (tamaño máximo de segmento).

Pero cuando lo importante en la comunicación es el uso de la memoria, se puede utilizar las siguientes recomendaciones para afinar o minimizar su uso.

- Si el tráfico es predominantemente local, utilizar paquetes del mayor tamaño posible, que sea soportado por el tipo de red. Esto minimizar la fragmentación de paquetes intercambiados por el sistema local. El costo de esta compensación es la fragmentación en *gateways* que conectan la red LAN con otras que utilizan un menor tamaño de paquete.

- Siempre que sea posible, los programas de aplicación deben de leer y escribir en cantidades de
 - Menos o igual a 512 bytes
 -
 - Ligeramente menos o igual a 4096 bytes (o múltiplos de esto).

3.7.3.3 Afinamiento de los adaptadores de colas de transmisión y de recepción

La mayoría de manejadores proveen un conjunto de parámetros de afinamiento del control de transmisión y recepción de recursos. Estos parámetros típicamente controlan las colas de transmisión y límites de recepción, pero además controlan el número y tamaño de *buffer* u otros recursos. Estos parámetros limitan el número de *buffer* para los paquetes que pueden ser transmitidos por la cola o limitan el número de *buffer* de recepción de paquetes. Pueden ser ajustados para asegurar que las colas son del tamaño suficiente para manejar las sobrecargas generadas por el sistema o la red.

Lineamientos generales

- Reajustar la transmisión de las colas cuando el procesador es mucho mayor que la red (más común en un sistema multiprocesador donde muchos procesadores transmiten a un único adaptador).
- Reajustar la transmisión de las colas cuando el tamaño del *buffer* del *socket* es grande.
- Reajustar la cola de recepción cuando el tráfico es muy pesado.

- Reajustar la cola de transmisión y recepción cuando existe una alta tasa de paquetes de tamaño pequeño.

3.7.3.3.1 Colas de transmisión

Para transmitir, los manejadores de dispositivos pueden proveer de un límite de cola de transmisión. Este puede ser a nivel de *hardware* o *software*, dependiendo del adaptador y del manejador. Algunos manejadores tienen únicamente una cola a nivel de *hardware*; algunos tienen ambos. Algunos manejadores internamente controlan las colas de *hardware* y únicamente permiten al *software* controlar los tamaños de la cola.

Generalmente los manejadores de los dispositivos van a transmitir directamente de la cola los paquetes, a la cola del adaptador del *hardware*. Si el procesador del sistema es relativamente rápido puede acelerar el tráfico en la red, o en un sistema SMP (multiprocesamiento simétrico), el sistema puede transmitir paquetes mucho más rápido de lo que son transmitidos en la red. Esto puede causar que la cola del hardware se llene.

Después de que la cola de hardware esté llena, algunos manejadores proveen una cola a nivel de *software* que empezarán a utilizar. Si la cola de transmisión de software también se llena, entonces los paquetes transmitidos son excluidos o rechazados. Esto puede afectar el rendimiento porque a nivel del protocolo debe de transmitir el paquete cuando no se confirma la recepción.

3.7.3.3.2 Colas de recepción

Algunos adaptadores permiten configurar el número de recursos a utilizar para la recepción de paquetes desde la red. Esto incluye el número de *buffer* de recepción (y sus respectivos tamaños) o puede ser el parámetro de la cola de recepción (el cual indirectamente controla el número de *buffer* de recepción).

Los recursos para la recepción de paquetes, pueden necesitar ser incrementados para manejar picos de sobrecarga. El manejador del dispositivo de interfase pone los paquetes en la cola de recepción. Si la cola de recepción está llena, los paquetes son eliminados y perdidos, como resultado el nodo debe retransmitir nuevamente los paquetes.

Tabla VIII. Parámetros para la evaluación y rendimiento de la red.

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIAS
Total de paquetes recibidos	Cantidad total de paquetes recibidos en un determinado periodo de muestras	Máximo posible	Indica que tanta actividad de trabajo es realizado por el sistema de la red
Número total de errores registrados	Proporciona la cantidad de errores por diversos motivos en la recepción de datos	Mínimo posible	Al ser un valor alto de la cantidad de errores, es necesario verificar la configuración y el tráfico de la red, para identificar los problemas
Porcentaje de los paquetes que son retransmitidos	Porcentaje de paquetes que son transmitidos nuevamente por varios factores.	Máximo 15%.	Cuando el número de paquetes es mayor a este porcentaje, existe una alta tasa de tráfico en la red. Es indispensable determinar si conviene mejorar los gateways y segmentar la red, para mejorar el tráfico.

Continuación 2/4

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIAS
Número de errores de salida	Proporciona una cantidad de errores de salida de datos o que no son recibidos o no llegan al destinatario de manera correcta	Mínimo posible	Cuando aumenta la cantidad de errores de salida, es probable que la información no llega a su destino o ésta no es aceptada por el destinatario por errores de recepción o fragmentación de la información, errores de chequeo al verificar la integridad de la información, o espacio insuficiente del <i>buffer</i> en el manejador del dispositivo de la red
Número total de paquetes transmitidos.	Cantidad de paquetes enviados al exterior.	Cantidad de paquetes transmitidos	Indicador de carga colocada en la red y es de utilidad para determinar el papel del nodo en la red.

Continuación 3/4

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIAS
Porcentaje de error en los paquetes recibidos	Es el porcentaje de los paquetes recibidos con errores	Máximo 1%	Cuando se tiene más de este porcentaje, verificar la disponibilidad de la memoria y la administración de las colas de recepción
Porcentaje de error en los paquetes enviados	Es el porcentaje de paquetes enviados y que presentan problemas en el envío	Máximo 1%	Cuando este porcentaje es mayor, es necesario verificar la administración de las colas de envío y de ser necesario aumentar el tamaño de estas colas
Tasa de colisión de paquetes	Tasa de colisiones de paquetes enviados en la red	Máximo 10%	Si el porcentaje de colisiones es mayor a este parámetro posiblemente sea necesario segmentar o reorganizar la red debido a una alta utilización de la misma

Continuación 4/4

PARAMETRO	DESCRIPCION	VALOR	SUGERENCIAS
Longitud de la cola de envío de paquetes	Contador que mide la longitud de la cola de salida de paquetes en espera a ser colocadas en la red	Máximo 2 unidades	Si el valor es mayor, es necesario mejorar los dispositivos o adaptadores por uno más recientes y más rápidos
Sobrecarga de la red.	Número máximo de errores por colisión más el número de transmisiones fallidas, dividido en la cantidad de paquetes transmitidos	Máximo 5%.	Si el porcentaje es mayor, probablemente sea necesario reorganizar la estructura de la red, para balancear la carga.
Fragmentación de paquetes.	Porcentaje de la cantidad de paquetes enviados por la red que son fragmentados.	Máximo 10%.	Cuando existe una fragmentación de paquetes, es posible que las colas de transmisión y recepción, tengan diferentes tamaños definidos o bien los gateway y nodos a través de la red, lo que es indispensable examinarlos y definir un valor constante, idealmente del tamaño del MTU.

4 CASO DE ESTUDIO DE EVALUACIÓN DEL RENDIMIENTO DE LOS SISTEMAS OPERATIVOS

En este capítulo final, se evaluará el rendimiento de los sistemas operativos con la ayuda de herramientas de monitoreo, las cuales proporcionan información fácil y oportuna para la identificación de las áreas críticas del sistema operativo. Se evaluarán los elementos asociados al rendimiento y se medirá los parámetros que deben ser modificados, sugerencias de valores que deberían de tener, etc., de manera que se pueda tener un mejor desempeño general.

Es necesario realizar un *backup* del sistema operativo previo a la evaluación y principalmente cuando se harán cambios en la configuración del sistema; en unos sistemas incluso permite realizar *backup* del *kernel*. Definir lineamientos y políticas sobre la aplicación de los cambios realizados, inclusive registrar los cambios en una bitácora o libro de cambios del sistema.

Con el fin de mejorar el rendimiento del sistema operativo es conveniente no realizar cambios a los parámetros y configuración general, sin un previo procedimiento de monitoreo para detectar el comportamiento de los componentes críticos, al realizar un monitoreo inicial se determina un punto de partida del comportamiento del sistema. Este punto inicial es indispensable para la identificación de los cambios necesarios a realizar y sirve de base para comparar que los cambios realizados mejoran el rendimiento o aún en el peor de los casos, se tenga un rendimiento más deficiente del anterior.

Otro punto importante es identificar los objetivos del afinamiento, generalmente se desea que todo quede mejor, pero esto en algunos casos es imposible, pues se puede lograr mejoras en algunas áreas, pero también a expensas de sacrificar otras áreas, y esto es de utilidad pues se puede orientar los esfuerzos en áreas que no son de interés para el negocio o bien afinar objetivos secundarios a expensas del objetivo primario, y aún más, incurriendo en costos y gastos no necesarios.

En este análisis de estudio, no se plantea un objetivo específico en el afinamiento del sistema operativo, debido a que la intención es determinar las áreas que muestran problemas y soluciones para mejorarlas, Además, no se puede implementar estas sugerencias, pues estas mediciones se realizaron en diferentes centros de cómputo en los cuales nos facilitaron el acceso para la evaluación, no así la facilidad de poder implementar los cambios y determinar la mejora producida por el cambio sugerido, esto debido a diferentes factores y motivos que están fuera del alcance de este trabajo.

4.1 Análisis de la información de los comandos de línea, con monitores gráficos

No es el objetivo dar a conocer una herramienta específica, o entrar en detalle en el uso de comandos propios de un sistema operativo en especial, por lo cual se detalla algunas figuras sobre el comportamiento de las áreas fundamentales de los sistemas operativos que son importantes en la evaluación y rendimiento. Se hace mención de los resultados desplegados y proporcionados por las herramientas de monitoreo en estas áreas para comprender los conceptos y aplicación de las variables que afectan estos componentes.

En sistemas operativos, que tienen como base un UNIX o similar, existen utilitarios para monitoreo desde la línea de comando, unos ejemplos de estos son: vmstat, times, sar, par, etc., que son comandos que proveen información valiosa acerca de lo que está pasando en el sistema.

Los sistemas operativos, como se describió en capítulos anteriores, tienen un número de contadores que miden internamente la actividad del sistema. Cada vez que una operación es ejecutada, un contador asociado es incrementado. Los monitores utilizan estos contadores para proporcionar un estatus de la actividad interna del sistema operativo.

En este ejemplo, se medirá el rendimiento de la memoria de un sistema Linux, usando el comando vmstat

```
$ vmstat 10 8640
```

El comando anterior tomará datos de un tiempo determinado el cual se indica con los parámetros en el comando, éstos indican lo siguientes

- 10 Es el intervalo entre cada muestra.
- 86 Es un día completo el cual es la cantidad de muestras a tomar.
- 40 $24 \text{ horas} \times 60 \text{ min} / \text{hora} \times 60 \text{ seg} / \text{min} = 86400 \text{ segundos}$, dividido el intervalo de segundos de cada muestra da como resultado 8640 muestras por un día de actividad.

El resultado de este muestreo es de la siguiente forma:

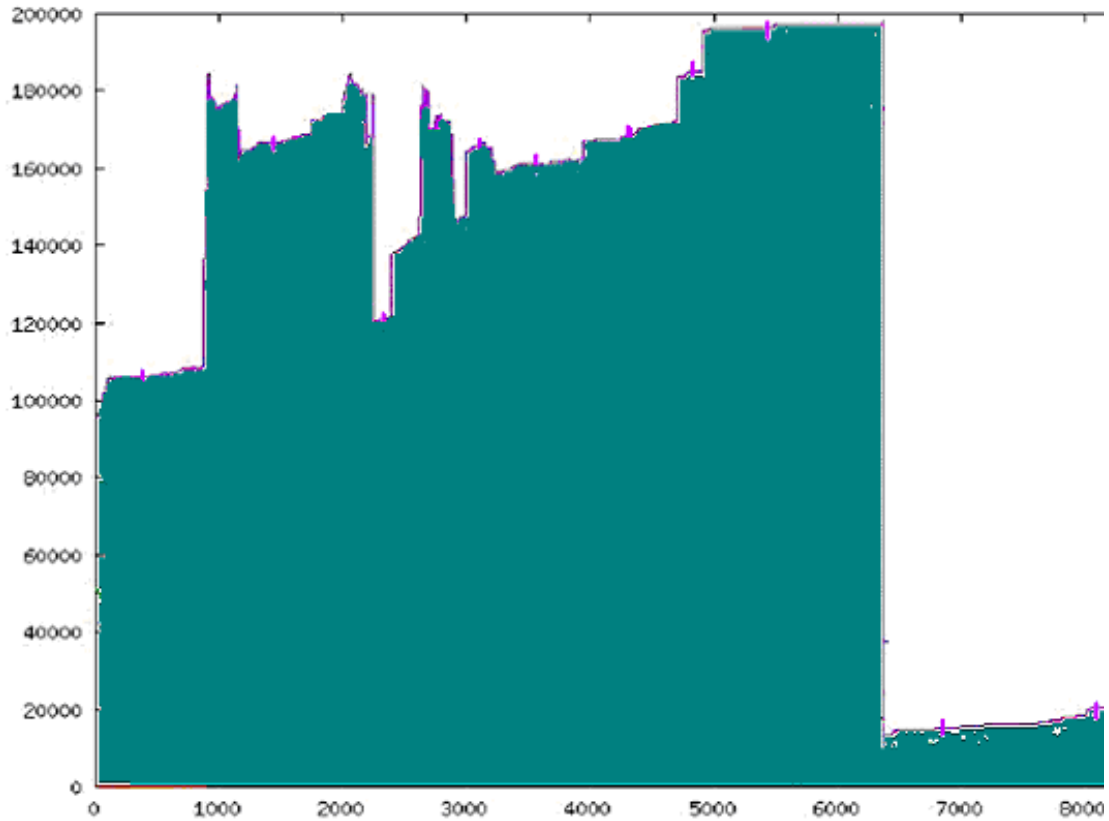
```
procs memory swap io system cpu r b w swpd free buff cache si so bi bo in cs us sy id
50840 42504 95704 0 0 1 1 124 32 0 0 100 0 0 0 0
50840 42504 95704 0 0 0 6 79 25 0 1 99 0 0 0 0
51436 42504 95708 0 0 0 8 225 267 3 1 96 0 0 0 0
51436 42504 95708 0 0 0 1 121 78 0 0 100 0 0 0 0
```

La información anterior es un poco difícil de interpretar, especialmente para personas que tienen poco conocimiento, especialmente en el uso de estos comandos y la información que proporcionan, lo interesante de las herramientas de monitoreo gráfico es que pueden leer la información generada por estos comandos de monitoreo y representar los datos de manera gráfica, de manera que es más fácil de interpretar y entender por parte de los administradores y usuarios en general.

Estos datos son leídos de diferentes formas, siendo la más común, por medio de archivos planos donde cada valor es separado por un carácter especial, ya sea una tabulación, una coma o bien un espacio en blanco.

Las siguientes gráficas son el resultado de la interpretación de una herramienta de monitoreo gráfico que lee la información obtenida por medio de comandos propios del sistema operativo, la información como se detalla anteriormente, es compartida por medio de archivos planos.

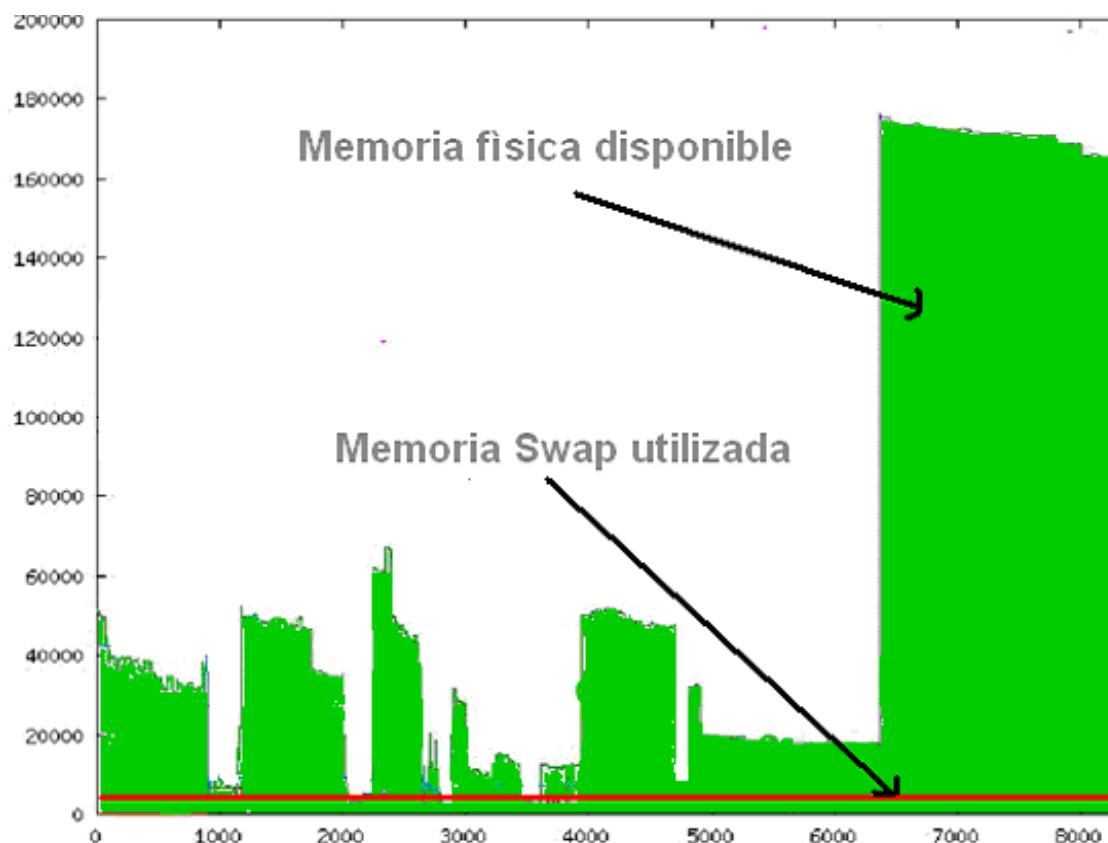
Figura 12. Comportamiento de la memoria caché



Comportamiento de la memoria caché, representado de manera gráfica medio de una herramienta que interpreta la información del comando VMSTAT.

Está figura denota una reducción de la disponibilidad de la memoria caché de manera drástica que tiene relación con la demanda de la memoria física en un determinado momento.

Figura 13. Rendimiento y disponibilidad de la memoria física y la memoria swap



Rendimiento y disponibilidad de la memoria física y memoria swap, de un servidor con sistema operativo Linux, en donde la demanda de memoria física minimiza la disponibilidad de la memoria caché, sin hacer uso de la memoria swap, que puede degradar el rendimiento.

Los datos tomados para estas figuras son información registrada del comando del sistema operativo Linux `$ vmstat 10 8640`, descrito anteriormente y de se puede hacer referencia al comportamiento del sistema.

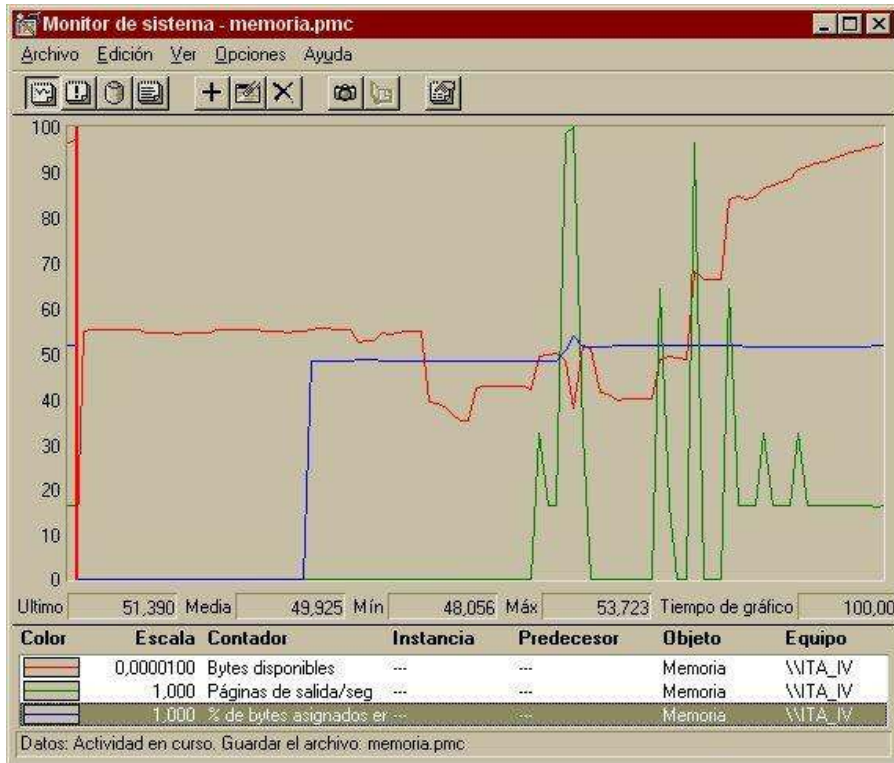
En las dos figuras anteriores, se puede interpretar de manera más fácil el resultado del comando `vmstat 10 8640.`, se puede observar que no hay grandes problemas de memoria, ya que la utilización del *swap* es bastante bajo, la cual nos da la idea de que el uso de la memoria no es crítico o bien hay suficiente memoria para los recursos y procesos existentes.

En la figura 12 se observa en el intervalo del eje X, entre 6000-7000 segundos, como en un instante la memoria caché baja y la cantidad de memoria disponible aumenta, con lo que se confirma que el rendimiento de memoria física es aceptable y sin problemas de rendimiento.

Los comandos proporcionados en sistemas operativos que tienen como base UNIX, son muy útiles y proporcionan gran cantidad de información de la cual se puede obtener estadísticas y datos del rendimiento de memoria, disco, procesador y red. Para éste estudio, se utiliza herramientas que representan los datos de manera gráfica, que son más fáciles de interpretar.

En el ambiente Windows®, se cuenta con herramientas de monitoreo que proporcionan el mismo tipo de información de manera gráfica, la cual es más entendible por el usuario final, tal como se muestra en la siguiente figura.

Figura 14. Monitores gráficos en ambiente Windows®

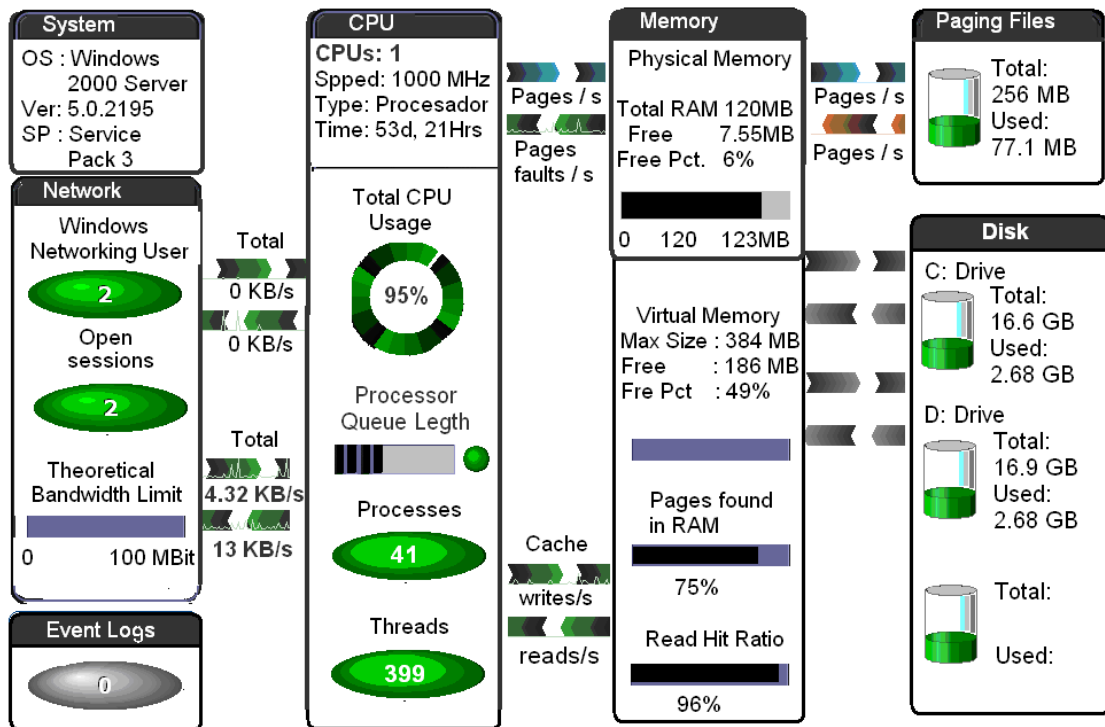


El monitor de sistema, permite determinar de manera más fácil la cantidad de memoria disponible, la cantidad de páginas escritas a disco y cualquier otro contador disponible con base en gráficas que describen el comportamiento de cada uno de los recursos.

4.2 Monitores gráficos para el análisis del sistema operativo

Como se ha mencionado constantemente, las herramientas de monitoreo gráfico, tienen la facilidad de proporcionar la mayor cantidad de información del sistema operativo, e indicar el estatus general del mismo en un determinado momento. En la figura siguiente se muestra los componentes de información del rendimiento y áreas de interés a evaluar del sistema operativo.

Figura 15. Herramienta de monitoreo para sistema operativo *Spotlight® on Windows*



Adaptado: Herramientas de sistema operativo *Quest Software*, <http://www.quest.com>

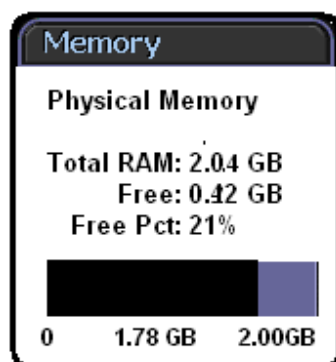
La información que proporciona esta herramienta, consiste principalmente en las características del sistema operativo que se está monitoreando, versión y parches instalados. Tráfico de red, uso del procesador, uso de la memoria y disco, como también de manera gráfica la cantidad de información y velocidad de información entre cada una de estas áreas.

La figura 15 de la herramienta de monitoreo gráfico, facilita a los usuarios administradores o bien encargados del rendimiento del sistema operativo, el estatus general del comportamiento del mismo, sin necesidad de recurrir a comandos en línea e interpretar información tabular que generalmente éstos proporcionan. No quiere decir que no es de utilidad el uso de comandos en línea propios de cada sistema operativo, si no la facilidad que la información gráfica proporciona en sí.

4.3 Análisis de la memoria

Para medir y mejorar el rendimiento del sistema operativo, que muestra ciertas dificultades de memoria se cuenta con la información siguiente

Figura 16. Medición de rendimiento y disponibilidad de la memoria



Esta pantalla gráfica proporciona información de la cantidad de la memoria física que dispone la computadora, que es de 2 GB, de la cual hay disponible entre 420 MB (0.42 GB) de memoria, que representa el 21% del total de memoria física.

Es recomendable que exista la mayor cantidad de la memoria física disponible, para que el sistema operativo en un momento determinado, pueda administrar memoria a un proceso urgente. El porcentaje ideal de memoria disponible es del 25% (75% de memoria puede estar comprometida o asignada).

4.3.1 Consecuencias de poca disponibilidad de memoria física

Cuando existe poca memoria física disponible en el sistema, las consecuencias pueden ser las siguientes

- Las aplicaciones existentes pueden ser incapaces de obtener más memoria.
- Nuevas aplicaciones pueden no ser ejecutadas por falta de memoria.
- El rendimiento en general es degradado por el uso excesivo de la paginación.

4.3.2 Aspectos a monitorear para determinar la poca disponibilidad y problemas de la memoria

Cuando se está ante una falta de memoria, es decir que está fuera de los límites permitidos o umbrales aceptables, es necesario verificar los siguientes componentes del sistema operativo.

- Verificar los procesos en ejecución, principalmente los datos sobre la cantidad de memoria física que están utilizando.
 - Cerrar las aplicaciones que hacen uso excesivo de memoria, y ejecutarse posteriormente o en otro momento determinado de ser posible.
 - Cambiar la cantidad de memoria a utilizar por aplicaciones, de ser posible por el sistema.
- Adquirir más memoria física de ser posible.

- Determinar si existe aplicaciones con mala administración de memoria o bien que no libera la memoria después de utilizarla.

Figura 17. Utilización de recursos de los procesos activos

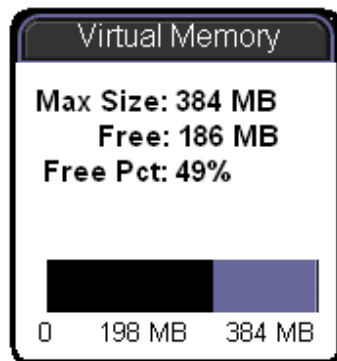
Processes								
Processes Services								
Process	PID	% CPU	Physic...	Virtual MB	Elapsed time	Handles	Threads	
Spotlight	1332	11.99	36.71	116.41	1:00	519	16	
WINWORD	1560	9.59	32.65	122.20	0:07	250	5	
ifbld60	1080	0.00	23.66	104.44	0:08	157	5	
TOAD	1444	0.00	21.55	75.96	0:59	145	4	
ifrun60	1512	0.00	15.00	79.21	1:00	102	4	
ifrun60	1516	0.00	13.61	75.08	0:40	100	4	
Babylon	1360	0.00	11.13	68.81	1:01	245	10	
inetinfo	820	0.00	9.40	51.78	1:02	559	27	
Rtvsan	640	0.09	8.92	65.63	1:02	219	33	
svchost	504	0.00	7.77	37.05	1:02	413	27	
services	216	0.18	6.27	33.05	1:02	576	30	
WMIProxy	860	0.00	5.87	34.42	1:00	142	6	
explorer	1136	0.98	5.58	73.99	1:02	383	14	
lsass	228	0.00	5.51	28.51	1:02	331	16	

Con esta información se puede determinar el consumo o utilización de recursos por cada uno de los procesos, es decir con exactitud quién consume más memoria (tanto memoria virtual como física), tiempo transcurrido desde la creación del proceso, *threads*, etc. El análisis de esta información es uno de los primeros pasos para la evaluación de rendimiento de uso de memoria.

4.3.2.1 Memoria adicional

De ser necesaria la adquisición de memoria adicional, es necesario determinar qué cantidad es recomendable adquirir, y para esto se utiliza los indicadores e índices de la utilización de memoria virtual.

Figura 18. Memoria virtual



Este indicador proporciona el tamaño total de la memoria virtual existente, esto es 384 MB de la cual 128 MB son de la memoria física y 256 MB por medio de disco. El tamaño adicional de memoria por disco debería de ser a lo sumo igual a la cantidad de memoria física, la cual es de 128 MB.

En total se tiene 198 MB de memoria virtual activa (384 MB de memoria virtual máximo, menos 186 MB de memoria disponible), la cual nos indica que se está utilizando 70 MB de memoria adicional, sobre la memoria física existente, por lo que de ser necesario la adquisición de más memoria, es recomendable que se adquieran por lo menos 128 MB adicionales, para tener un total de 256MB de memoria física¹¹.

Otro factor en la medición de la evaluación de la memoria es la cantidad de paginación y *swap*, que son índices estrechamente relacionados con el rendimiento del disco, pero que dan referencias a la utilización de la memoria, en términos de cantidad de uso. También está el índice de la cantidad de aciertos de la información disponible en memoria, el cual indica como están definidos los parámetros sobre bloques de memoria para archivos y datos.

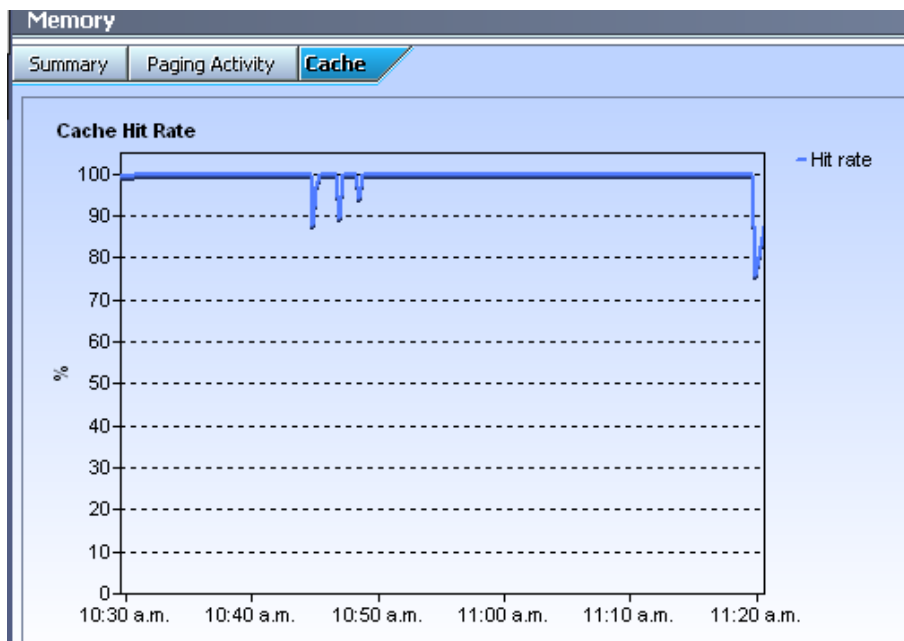
¹¹ Las cantidades de memoria aquí son pequeñas, pero para un servidor se espera que sea mayor la cantidad de memoria, pero el análisis para definir la cantidad de memoria adicional necesaria es el mismo.

4.3.2.2 Memoria caché

La memoria caché, generalmente es utilizada para mejorar la comunicación entre procesador-memoria-disco. Estas medidas de memoria caché, se indican generalmente en cantidad de páginas por segundo.

Un alto porcentaje de lectura de caché, no necesariamente indica un problema en el sistema operativo, cuando esto ocurre es necesario examinar los procesos existentes y ver cuales son los que tienen una mayor cantidad de carga.

Figura 19. Eficiencia en la utilización de la memoria caché

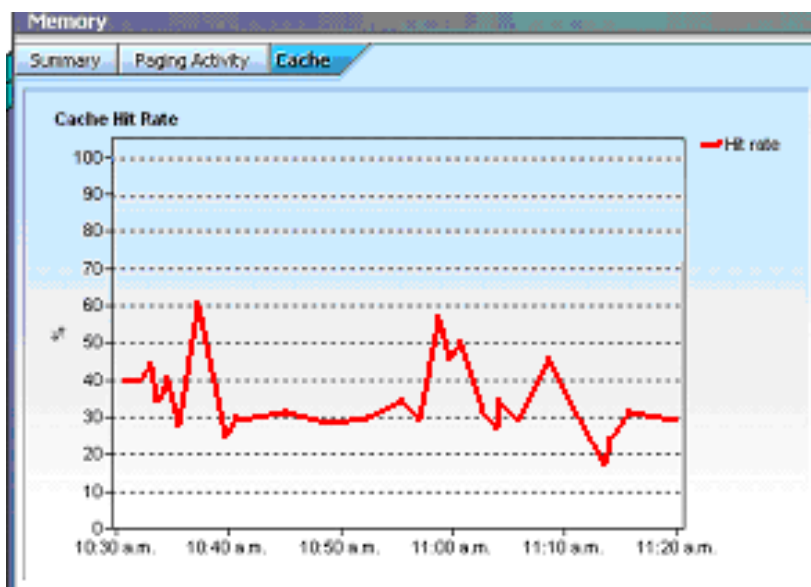


Idealmente se espera que toda la información solicitada se encuentre previamente en la memoria, lo cual permite un mejor rendimiento general del sistema operativo, el inconveniente es que esta memoria no puede ser demasiado grande, porque limita la cantidad de memoria a utilizar por las aplicaciones y si es demasiada pequeña, constantemente estaría poniendo y quitando bloques de datos.

La tasa de éxito de la memoria caché se espera que sea de por lo menos el 80% de las búsquedas, en este caso se tiene casi el 100% de éxito, con leves instantes de disminución, que resulta ser normal y aceptable.

En la siguiente muestra, se observa una tasa de aciertos de caché muy baja y constante (aproximadamente entre 30 y 40 por ciento de acierto), lo cual significa la existencia de una gran actividad de lectura de información.

Figura 20. Deficiencia de rendimiento de la memoria caché



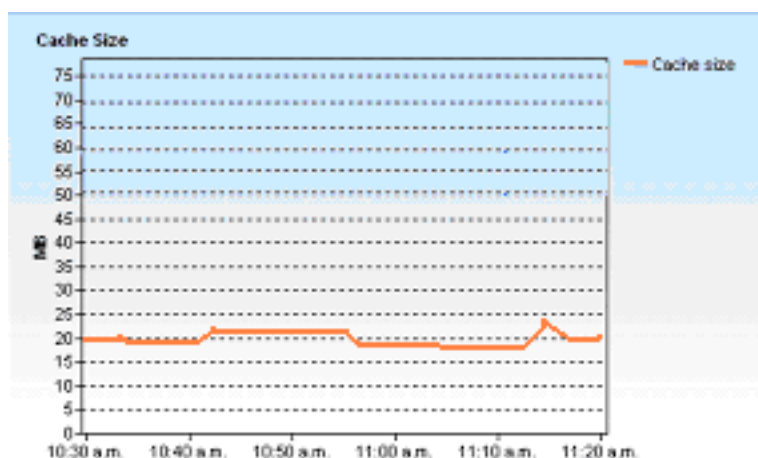
Esta ineficiencia del caché puede ser debido a dos factores principales que a continuación se describe.

- El rendimiento de acceso de la información a disco es ineficiente, debido a que utiliza un recurso compartido en la red, tal como un servidor NFS.
- El tamaño utilizado para la memoria caché es pequeño en relación a la cantidad de memoria del sistema. Este tamaño debería de ser aproximadamente entre un 15 y 20 % de la memoria física.

En este caso el acceso o rendimiento es aceptable, no muestra problemas de disco, además de ser recursos locales y utiliza una arquitectura *RAID* tipo 5.

El otro punto es el tamaño de memoria asignada a caché, que se muestra a continuación.

Figura 21. Indicador del tamaño de memoria caché utilizada



El tamaño aquí representado, está aproximadamente entre 20MB, que equivale aproximadamente a un 6%, del total de memoria (384MB)¹². Por el bajo rendimiento de la memoria, sería conveniente ampliar a unos 50MB a 60MB el tamaño de la memoria caché (esto es 15% de memoria física), para mejorar la tasa de éxito de datos en memoria.

4.3.2.3 Paginación de la memoria

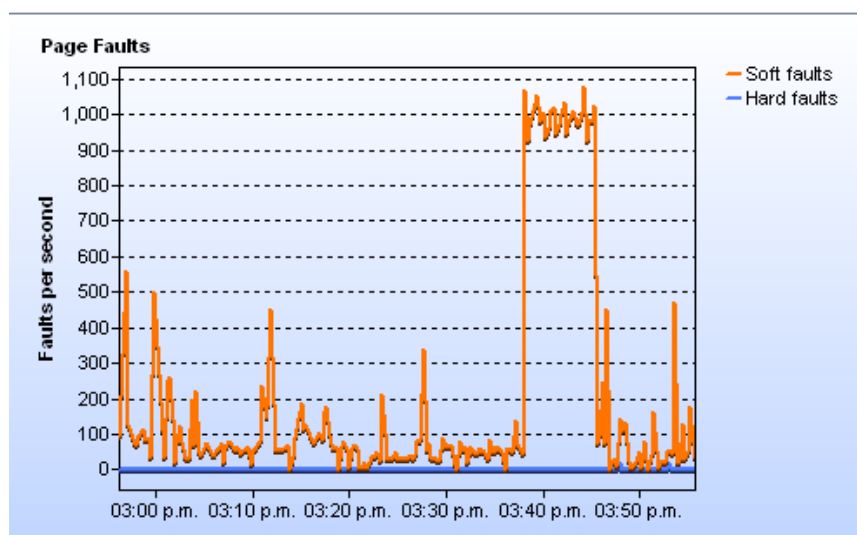
Otros factores para la medición del rendimiento de la memoria, corresponden a la cantidad de páginas de memoria que son movidas hacia y fuera de la memoria principal.

Cuando se solicita una gran cantidad de datos de memoria, puede ser que éstos no se encuentren disponibles y tenga que leer los datos directamente del disco. Cuando los datos se encuentran en memoria, pero fuera del área de trabajo del proceso que lo solicita, se denomina *soft faults*, y cuando es necesario recurrir al disco se denomina *hard faults*.

La siguiente figura se puede ver la cantidad de páginas solicitadas por los procesos, este indicador generalmente es bajo, con excepción cuando la carga es intensa. En este caso en particular, se puede observar que las páginas solicitadas generalmente se encuentran en la memoria, pues el indicador de páginas buscadas en el disco *hard faults*, es cero o uno.

¹² Estos datos fueron tomados de diferentes lugares y muestras, por lo que las cantidades aquí descritas varían, tomando como referencia la muestra que mejor se adapta para la explicación del tema.

Figura 22. Páginas de memoria solicitadas que no se encuentran disponibles en el área de trabajo



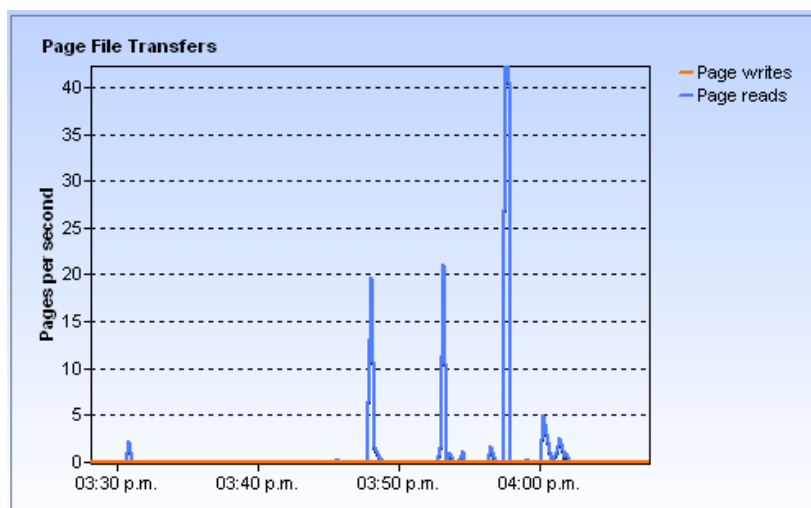
Cuando la actividad de paginación es alta y tiene que recurrir al disco para recuperar la información deseada, es recomendable aumentar la memoria física. En el caso de disponer de suficiente memoria, es posible que sea necesario modificar los parámetros de porcentaje de área reservada para páginas de datos y páginas de archivos o programas, según la orientación del servidor.

Cuando el servidor es para almacenamiento de archivos o programas, es conveniente tener un 66% de asignación para páginas de archivos y un 34% de asignación para páginas de datos, del total de memoria reservada para paginación.

4.3.2.4 Transferencia de páginas de archivo

Este indicador de transferencia de páginas de archivo, indica la cantidad de páginas que han sido escritas en el disco (*page writes*) y la cantidad de páginas que han sido leídas desde el disco (*page reads*), para lo cual un valor aceptable es de 5 páginas por segundo, y como se puede observar, esta es una tendencia aceptable como se ve en la gráfica siguiente, con excepciones en determinado momento en las páginas leídas.

Figura 23. Transferencia de páginas de archivos del disco a la memoria



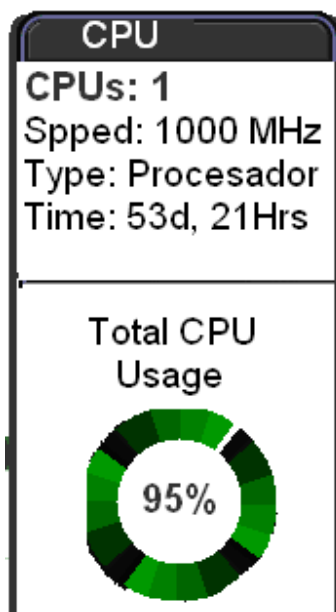
Una característica importante es el tiempo de transferencia de las páginas de datos desde y hacia el disco, que no debe sobrepasar de 10 milisegundos por transferencia. Como se observa, muchos datos proporcionan información, lo cual permite entender el desempeño de más de un área, como en estos casos del rendimiento de la memoria y del disco.

4.4 Análisis del procesador o CPU

El procesador o CPU, es un recurso importante en el rendimiento, que generalmente, está estrechamente relacionado con otros recursos como memoria, disco, etc.

Aquí se analiza algunos aspectos directos con problemas del procesador, que consiste en identificar los procesos y las consecuencias del bajo desempeño, debido principalmente a una sobrecarga de tareas por realizar.

Figura 24. Uso del procesador



En esta muestra de la herramienta de monitoreo, se puede ver información del procesador, el cual indica la cantidad de procesadores existentes, la velocidad de operación, tipo de procesador y tiempo de operación del sistema. Como también proporciona información de uso del procesador, el cual idealmente se espera con carga normal de procesamiento entre un 80%. Con valores más altos, es probable que en cargas altas el procesador no pueda atender todas las solicitudes y envíe los procesos a la cola por un tiempo considerable.

Cuando el índice de trabajo del procesador está por arriba del límite normal de procesamiento por un tiempo considerable, es necesario examinar qué esté provocando una utilización continua del procesador.

4.4.1 Consecuencias sobre problemas del procesador o CPU

En la evaluación del procesador, se debe de tomar en cuenta dos aspectos a, los cuales se derivan en el bajo rendimiento y la sobre utilización o demanda excesiva de uso del procesador.

Estos aspectos que hacen ver el rendimiento del procesador como un problema que afecta el rendimiento del sistema operativo en general, tiene repercusión en las áreas necesarias para evaluar y corregir los problemas. Sin embargo, el uso excesivo del procesador aún cuando no sea realmente un problema en el rendimiento, si hace pensar que en un determinado momento, éste puede ser el causante de un bajo desempeño del sistema en general, por la sencilla razón de que se limita la cantidad de trabajos a realizar, porque no se puede atender más o no se pueden atender de manera rápida.

No importando los aspectos que derivan la sospecha del bajo rendimiento, se deben de tomar en cuenta los siguientes lineamientos, para la revisión de uso prolongado de procesador:

- Observar los procesos y listarlos por consumo del procesador, de ser necesario eliminar los procesos que no responden en el sistema, o bien un proceso esté realizando una actividad inesperada.
- Considerar actualizar el procesador por uno más rápido o adicionar más procesadores al sistema.
- Observar la actividad de paginación de memoria, en busca de un proceso que tenga una alta actividad que pueda aumentar la utilización de. procesador.

4.4.1.1 Bajo rendimiento del procesador

Se debe de tomar en cuenta que el bajo rendimiento del procesador, puede ser consecuencia de otro recurso con el cual el procesador esté asociado o exista un enlace para realizar sus funciones, como podría ser la asignación de la memoria o solicitudes de páginas de datos o archivos del disco, las cuales cuando tienen un mal desempeño, afectan al procesador, indicando que tiene un bajo rendimiento por espera de otro recurso.

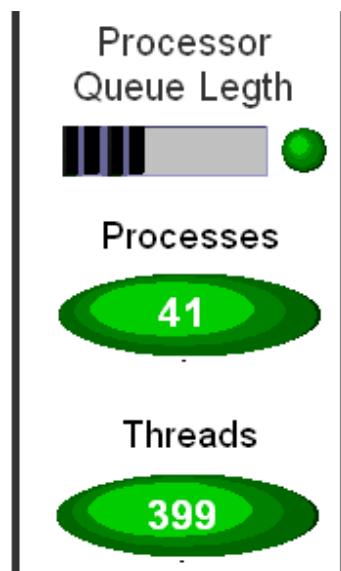
4.4.1.2 Demanda excesiva de uso del procesador

La demanda excesiva de uso del procesador, se enfoca principalmente cuando existen muchos procesos o servicios que mantienen ocupado al mismo y llega a determinado punto en el cual éste ya no es capaz de atender a todos, de manera pronta y oportuna.

4.4.2 Análisis de procesos y *threads* existentes

Los procesos en ejecución en el sistema se muestran en la siguiente figura, donde se puede observar como información la cantidad de procesos existente y *threads*, que compiten constantemente por el uso del procesador.

Figura 25. Cola de procesos y cantidad de procesos y *threads* en el sistema.



En sistemas de alto desempeño, es normal que puedan tener hasta 10 ó 15 procesos en la cola de espera como normal y con un desempeño aceptable, pero en otros sistemas 5 procesos en la cola de espera es considerado como una cantidad alta en espera por el procesador¹³.

4.4.2.1 Cola de procesos

En la figura anterior, se puede ver que existen aproximadamente 4 *threads* en espera de asignación del procesador, generalmente cuando la longitud de la cola de procesos en espera está en los umbrales o límites máximos, se debe a que el procesador no dispone de la capacidad para atender todas las tareas y de una manera sustancial, afecta el rendimiento general del sistema.

¹³ En ambiente Windows® 5 procesos en cola por procesador es ya un indicador de sobrecarga.

Cuando se tiene mucha actividad que el procesador no puede atender, se recomienda realizar algunas de las siguientes opciones

- Determinar los procesos con mucha actividad en el procesador.
- Analizar la posibilidad de mover algunos procesos o actividades a otro sistema.
- Adicionar más procesadores al sistema, los *threads* en espera serán distribuidos entre los procesadores existentes.
- Actualizar el procesador por uno más rápido, esto permitirá que los procesos sean atendidos antes, pero posiblemente no tengan el mismo desempeño que al adicionar más procesadores.

Con esta cola de procesos lo que se puede determinar es que efectivamente el procesador tiene problemas de rendimiento, lo cual se refleja en la cantidad de procesos en espera a ser atendidos en un momento determinado. Claro está es necesario tomar varias muestras y si la tendencia se mantiene, se puede decir que este es un indicio de una degradación generalizada del sistema como consecuencia del rendimiento del procesador.

4.4.2.2 Distribución de actividades por el procesador

Normalmente, la distribución del tiempo de servicio del procesador en tareas propias del usuario, tareas del sistema e interrupciones y espera del disco es una relación 60-30-10, la cual es considerada como una medida aceptable como medida de referencia.

Cuando la actividad por el sistema aumenta, se considera que existe una demora en los servicios propios del sistema, como una alta llamada de servicios de interrupción o tiempos de atención de las interrupciones demasiadas altas, las cuales deberían de verificarse y determinar según especificaciones las medidas descritas por los fabricantes.

Figura 26. Porcentaje de atención de diferentes actividades del procesador



En la figura 26, el procesador atiende actividades propias del usuario en un 61%, actividades propias del sistema 32% (éstas incluyen interrupciones, llamadas al Kernel del sistema operativo, etc.), y por último tiene un 7% por espera de acceso o restricción de disponibilidad del disco.

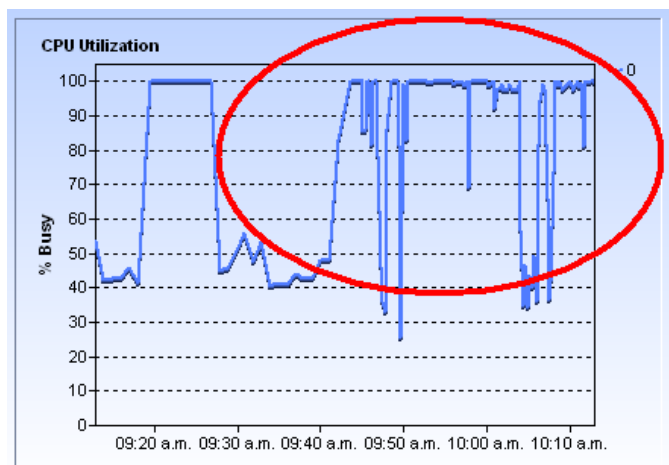
Esta muestra se encuentran en los límites aceptables, en donde se puede observar principalmente que los atrasos en el procesador no es por espera de disco en este momento, por lo que es necesario analizar los procesos en ejecución.

4.4.2.3 Utilización del procesador

El procesador detalla la cantidad de trabajo realizado en un momento determinado, el cual siempre realiza más de alguna modificación aún cuando aparentemente no se realice una actividad o no se tenga sesiones activas.

Durante las actividades normales de uso del procesador, se tiene un promedio de 60% de actividad, arriba de este indicador, es considerado como un porcentaje alto de actividad hasta un 90%, en el cual se considera una actividad alta de uso, que debe de ser analizada para determinar que hace que el procesador esté muy atareado.

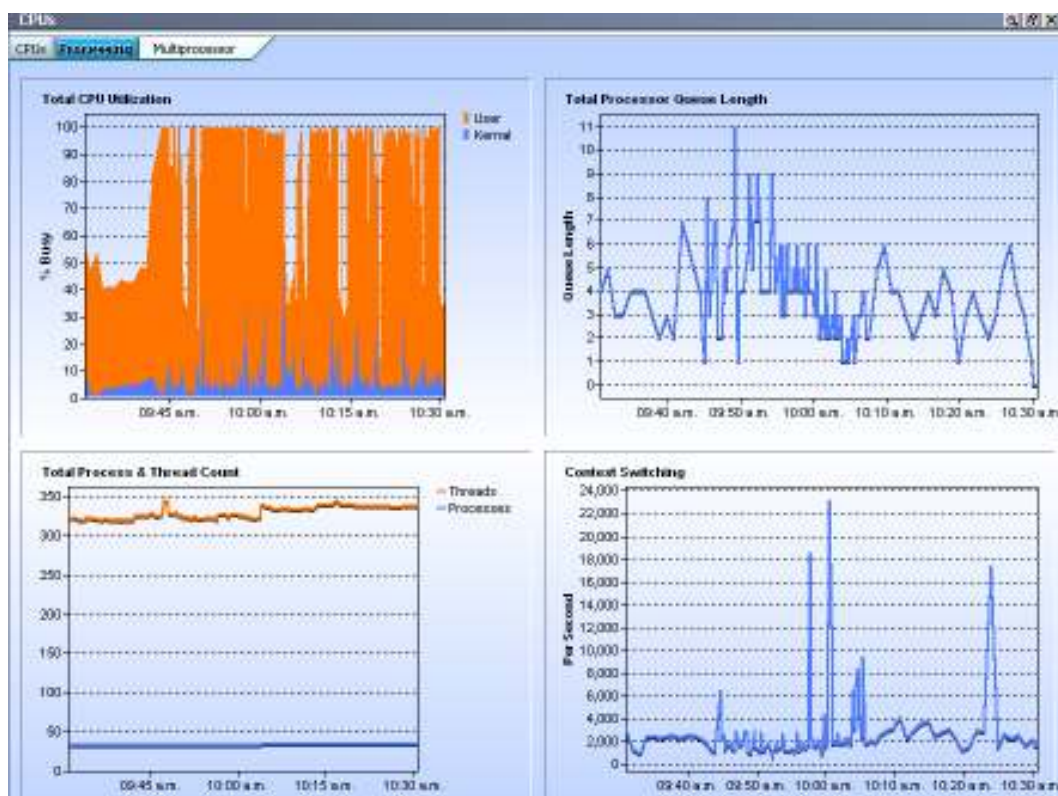
Figura 27. Utilización del procesador



En la gráfica anterior, el procesador está ocupado casi al 100% (o cerca de este intervalo, tal como se muestra en el círculo sobre esta figura), en la mayoría de tiempo, que es constante, a excepción de instantes determinados, en la cual tienen bajas por unos segundos.

Determinar la alta utilización del procesador, lleva al siguiente paso, el cual es determinar los procesos con mayor carga y poder determinar si es posible mejorar el rendimiento.

Figura 28. Actividad de procesamiento atendido por el procesador



Total de utilización de CPU, Cantidad de procesos y *threads* existentes, longitud de cola del procesador y cantidad de cambios de contextos por unidad de segundos.

Como se puede confirmar, el procesador tiene una carga constante de mucha actividad, se tienen varios procesos en espera de manera continua alrededor de 5 a 8 procesos en promedio y la cantidad de contextos en determinados momentos es alta en comparación al ritmo normal.

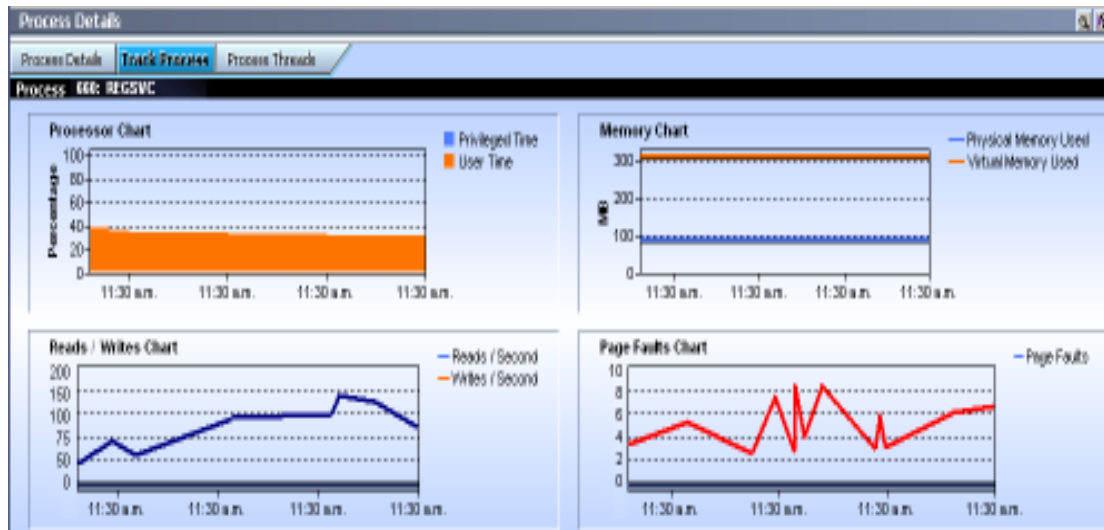
Figura 29. Estadísticas de procesos en ejecución



Esta figura proporciona la información de los procesos en ejecución, detallada en el círculo, dando información de la cantidad de tiempo asignado del procesador, memoria física asignada y otras estadísticas interesantes.

Se pueden ver más a detalle valores y parámetros relacionados en el rendimiento y ejecución de los procesos, en la parte de abajo, se describen los detalles de un proceso específico llamado regsvc.

Figura 30. Seguimiento de un proceso determinado



Observación del porcentaje de uso del procesador, cantidad de memoria utilizada en general, índices de lecturas y escrituras realizadas por segundo, como el índice de páginas fuera del contexto.

En este detalle se puede ver la actividad del proceso REGSVC, el cual tiene una cantidad considerable de bloques de memoria para lectura y escritura, y tal como describe en el cuadro inferior izquierdo, en la opción de *reads/writes charts*.

Otro punto interesante es la cantidad de páginas de memoria que solicita y no se encuentran disponibles, las cuales están entre 4 y 8 páginas por segundo, y se puede decir que en promedio es de 6 páginas por segundo, el cual es un valor alto, ya que generalmente un promedio de 5 páginas por segundo, es considerado en una carga alta.

Un punto interesante es que el sistema reporta en la distribución de utilización del procesador, que el tiempo dedicado por operaciones de disco es bajo. Esto lleva a la conclusión de que el principal problema de los procesos es la disponibilidad de memoria para la asignación de páginas de información, por lo que de ser posible es se debe de asignar más memoria física al sistema.

4.4.2.4 Carga de los procesos

La carga de procesos en un mismo sistema operativo es también la causante del bajo rendimiento del sistema, esto es debido a la cantidad de procesos existentes y a la cantidad de requerimiento de asignación del procesador.

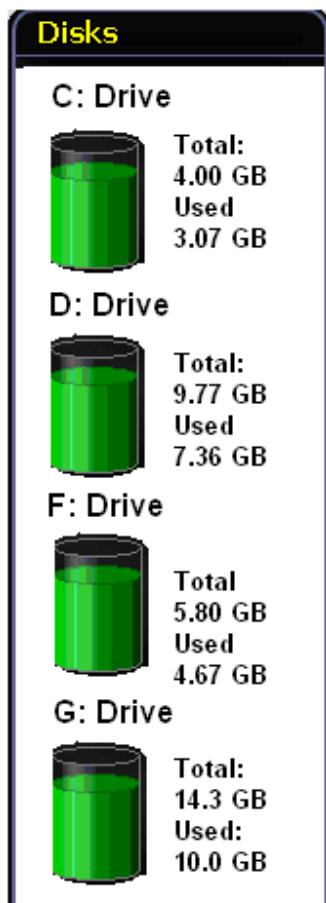
Como se puede observar en la figura 30 sobre las estadísticas de los procesos en ejecución, existen muchos procesos y las necesidades de asignación del procesador son altas. Es conveniente determinar si alguno de los procesos listados, puede ser cambiado o movido a otro servidor para su ejecución, o bien de ser posible, ejecutar el proceso en un horario diferente, pues el procesador no tiene la capacidad suficiente para dar un buen desempeño dada la cantidad y carga de trabajo.

La solución de ejecución de carga de trabajo en horario diferente generalmente es una solución factible y a corto plazo, el principal inconveniente que se encuentra en estas soluciones es el factor humano que necesita estar presente, cuando el proceso requiere de confirmación o ingreso de un dato específico durante su desarrollo o ejecución.

4.5 Análisis del disco

El disco es parte de los recursos que tiene impacto sobre el rendimiento general y debe de ser estructurado desde el inicio de la implementación del sistema operativo. Para este recurso es mejor planificar su utilización y capacidad de crecimiento al inicio de la configuración del sistema, aunque siempre es posible mejorar el rendimiento.

Figura 31. Unidades lógicas del disco en un sistema operativo



Este sistema cuenta con cuatro unidades lógicas denominadas con las letras C, D, F, G.

En cada una de ellas se detalla la capacidad y el tamaño utilizado.

En este sistema se tiene la unidad C, la cual tiene las aplicaciones propias del sistema operativo y un área temporal.

El drive D y F, son utilizados para almacenamiento de datos y un área temporal.

La unidad G es asignada para aplicaciones y como un repositorio de archivos que son compartidos por los usuarios a través de la red.

4.5.1 Consecuencias de un mal rendimiento del disco

El sistema operativo utiliza discos para almacenar datos e información relacionada con los procesos en ejecución. El proceso de lectura y escritura consume una cantidad considerable de recursos, por lo que es necesario tener un óptimo desempeño y eficiencia cuando se realizan operaciones de lectura y escritura.

El sistema de almacenamiento también es utilizado para grabar temporalmente una porción de la memoria, que permite hacer referencia a una mayor cantidad de memoria física, esto implica que el área de memoria asignada para este proceso debe de ser lo más eficiente posible, minimizando los retardos a la información almacenada y al mismo tiempo ser utilizada lo menos posible, es decir, que la memoria física debe de ser lo suficientemente grande para no necesitar frecuentemente el uso de este tipo de memoria al disco, que recibe también el nombre de área de paginación.

Otro uso que se realiza con los discos es para *swap*, que como se ha indicado en capítulos anteriores, permite intercambiar datos almacenados en la memoria física con datos que residen en el disco.

Se debe tomar en cuenta que el procesador generalmente trabaja con datos, es decir realiza operaciones de cálculo y procesamiento y éstos los lee de memoria y esté a la vez los solicita al disco cuando no los tiene presente, lo que hace pensar que puede afectar el rendimiento del procesador.

4.5.2 Aspectos a monitorear por el bajo rendimiento del disco

Para aplicaciones que tengan una gran carga de procesamiento, o simplemente medir el rendimiento del sistema de discos, es importante revisar los siguientes parámetros.

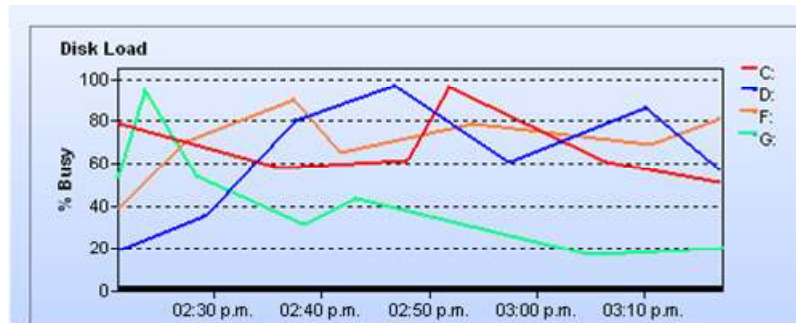
- Porcentaje de disco ocupado
- Porcentaje de lectura de caché del disco
- Porcentaje de escrituras de caché del disco
- Porcentaje de tiempo de espera por el disco

Estos indicadores son ideales para determinar el rendimiento y sugerencias a mejorar en el sistema.

4.5.2.1 Factor de utilización de los discos

En las mediciones sobre utilización del disco, se tiene en promedio que éste se encuentra en un factor de 70 y 80 por ciento ocupado, y en cierto momento aumentan la actividad alrededor del 90%, que posiblemente se deba a la cantidad de carga. El valor aceptable de utilización de un disco debe de ser menor del 60% en cargas normales de actividad, por arriba de este porcentaje, el factor de utilización es alto, por arriba del 90 al 95 por ciento la carga de utilización de disco ya presenta problemas por sobrecarga, lo que sería ideal mejorar el rendimiento y utilización de estos dispositivos.

Figura 32. Factor de utilización del disco



El disco G, donde se almacena parte de las aplicaciones y textos compartidos, se utiliza menos. Éste detalla una alta actividad al inicio del periodo de muestra y conforme pasa el tiempo, la actividad en este dispositivo es menor, esto generalmente como resultado del comportamiento del sistema, a la carga de los usuarios de sus aplicaciones al inicio de la jornada de trabajo.

En disco C, donde se encuentra instalado el sistema operativo y área temporal de datos, el factor de utilización es alto, por lo que se tendría que realizar un mayor análisis, debido a que este es el principal medio que utiliza el sistema operativo para el almacenamiento de páginas de memoria y área de *swap*.

Figura 33. Resumen de actividad de los discos.



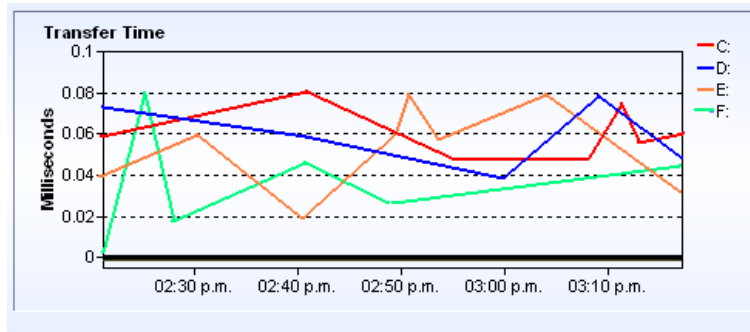
En la figura anterior de la actividad del disco, se puede observar que los discos suman un 81% de actividad entre lectura y escritura y el resto, que corresponde a 19% es de inactividad del disco.

4.5.2.2 Transferencia del disco

El tiempo de transferencia de información del disco, es un factor importante para determinar que tan eficiente es el disco para la carga de trabajo desempeñado.

Constantemente estos factores cada vez son menores por la tecnología implementada.

Figura 34. Tiempo de transferencia de los discos



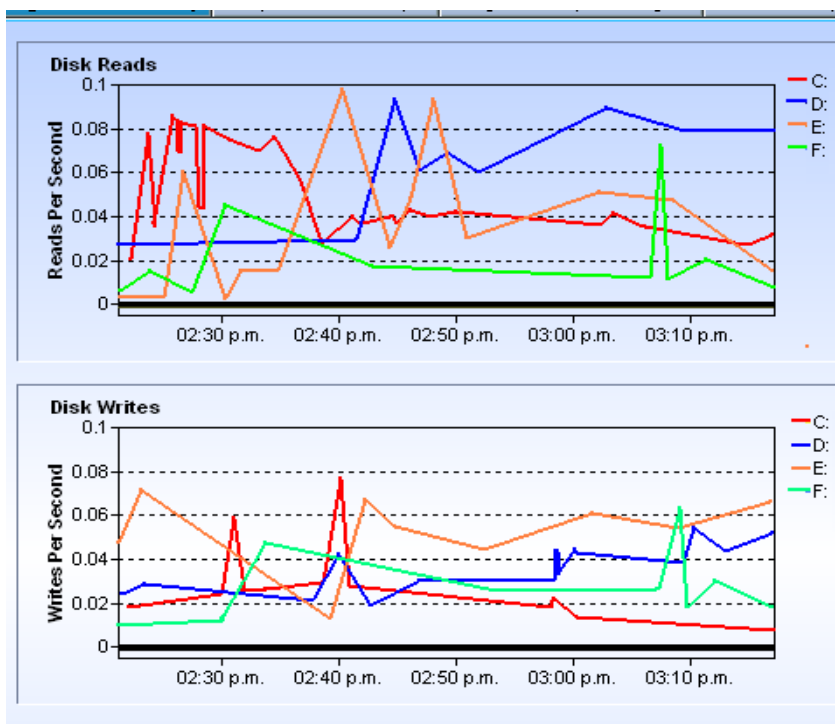
Estos tiempos oscilan entre 0.02 y 0.08 milisegundos por transferencia, los cuales tiene un valor aceptable considerando la carga y factor de utilización del disco. Los cuales son factores aceptables, que están por debajo de los 10 milisegundos, que es una base ideal del rendimiento de disco.

4.5.2.3 Lecturas y escrituras del disco

La cantidad de lecturas y escrituras al disco, es un indicador complementario para determinar la eficiencia del disco con base en el porcentaje de tiempo de disponibilidad y utilización del disco.

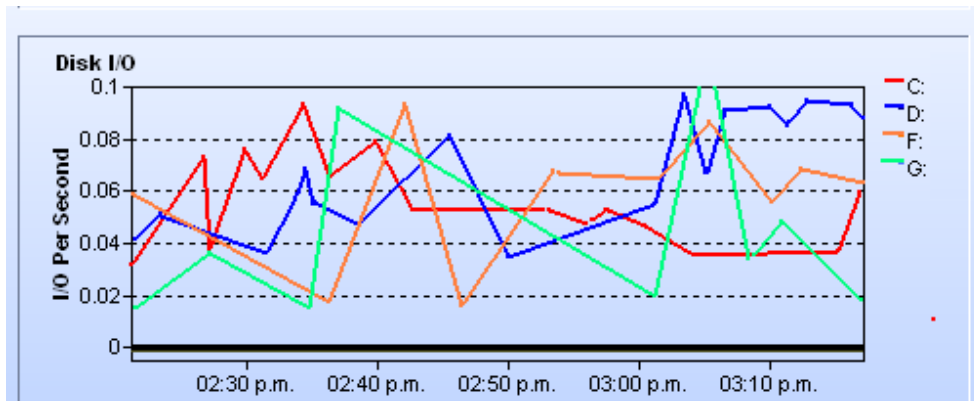
Estos factores son determinantes para saber qué tan bien responden los discos al momento de transferir información a la memoria.

Figura 35. Escritura y lectura de discos en un rango de tiempo determinado.



En esta gráfica se puede ver el comportamiento individual de cada disco con respecto a las escrituras y lecturas por segundos.

Figura 36. Actividad de I/O reportado por segundo de disco.



En la figura anterior se puede ver la actividad de utilización de los discos para lectura/escritura por segundos. Estos indicadores dan una idea sobre la actividad que realiza de manera generalizada cada uno de los componentes en la transferencia desde y hacia los discos.

Es importante tomar en cuenta que el porcentaje de utilización del disco entre lectura, escritura y disponibilidad de disco no es mas del 100% de actividad, y de manera generalizada la actividad para los cuadro discos en el periodo determinado está entre un 81% de lecturas y escrituras.

4.5.2.4 Tiempo de espera por disco

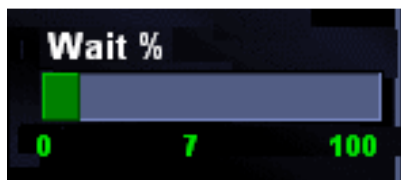
El tiempo de espera por disco es un factor importante que también es de gran utilidad para el desempeño del procesador, pues generalmente el procesador espera un tiempo determinado por datos para procesamiento cuando son requeridos directamente del disco.

Los motivos por los cuales el disco tiene un desempeño inferior, que hace esperar principalmente al procesador y a otros recursos son debido a factores mecánicos involucrados al hacer referencia a la información almacenada, tales como los brazos de lectura y escritura, o bien como la cantidad de revoluciones del disco.

Un punto interesante para mejorar el desempeño es tener una baja fragmentación de la información, especialmente cuando se realiza mucha lectura de manera secuencial, lo cual evita el movimiento continuo de los brazos de lectura y escritura.

La figura siguiente, detalla la actividad del procesador, en ésta se observar que en un momento determinado, el procesador espera por actividades de disco un 7%.

Figura 37. Espera de procesador por el disco

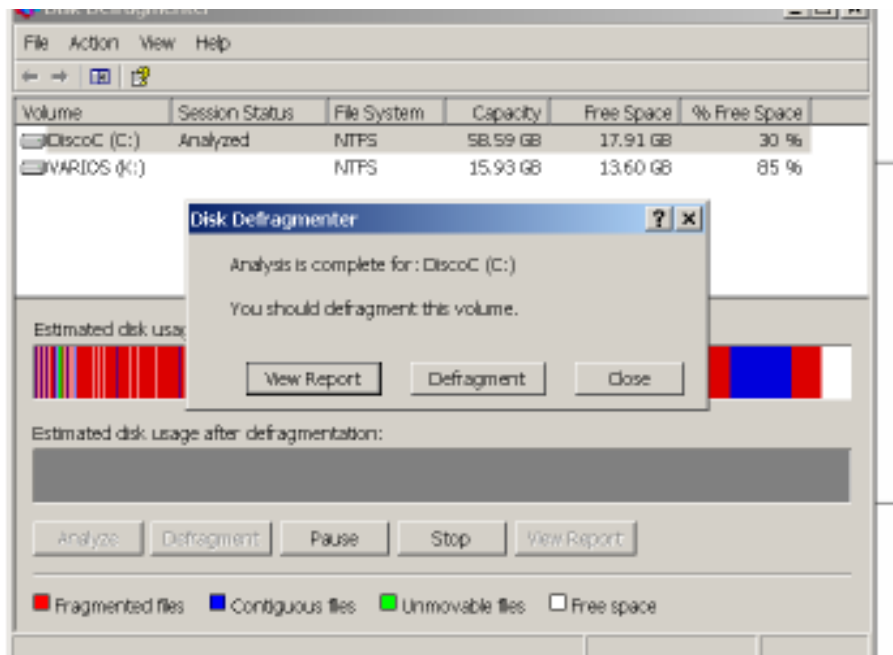


Los tiempos de espera para lectura y escritura del disco por parte del procesador, no deben de ser por arriba del 20%, del total de actividad del disco.

4.5.2.5 Fragmentación del disco

La fragmentación es otro factor importante en la utilización de los discos, pues esto nos indica que tan fácil es el acceso de los datos de manera secuencial.

Figura 38. Medición de la fragmentación, utilizando el *disk defragmenter* de Windows®



En esta muestra, se tomó con *disk defragmenter* de Windows®, el cual indica de manera gráfica el porcentaje de archivos fragmentados en color rojo y sugiere realizar una de fragmentación del nuestro disco.

Este tipo de análisis es necesario realizarlo con frecuencia para tener un disco en óptimas condiciones y mejorar los tiempos de acceso. Generalmente es aceptable un 15% de fragmentación, por arriba de un 20% es necesario reducir la fragmentación.

4.5.2.6 Resumen del rendimiento del disco

Como se puede observar de manera generalizada, el sistema de discos a pesar de que tiene una alta carga del trabajo, por la cantidad de actividad constante que realiza y los índices de lectura y escritura que son altos, presenta un desempeño aceptable, lo cual es representado por el tiempo que hace esperar al procesador.

Se puede tratar de agilizar un poco la actividad de los discos, de manera que se pueda distribuir la carga. Una manera interesante de aprovechar los discos en su aspecto del factor de utilización y mejorar la cola de espera, es poder distribuir la información en estos cuatro discos, esto idealmente por medio de una configuración de disco RAID, con la cual se tendría un mejor desempeño general.

La partición es otro punto interesante para el mejoramiento del rendimiento del disco, esto es determinar las aplicaciones más utilizadas y de ser posible ubicarlas en otro sistema disponible. Pero como se mencionó anteriormente, el rendimiento es aceptable desde el punto de vista de la atención del procesador.

4.5.3 Embotellamiento por utilización de los discos

Para la identificación de embotellamiento del disco es necesario medir los siguientes índices.

- Porcentaje de utilización del disco
- Porcentaje de disponibilidad del disco

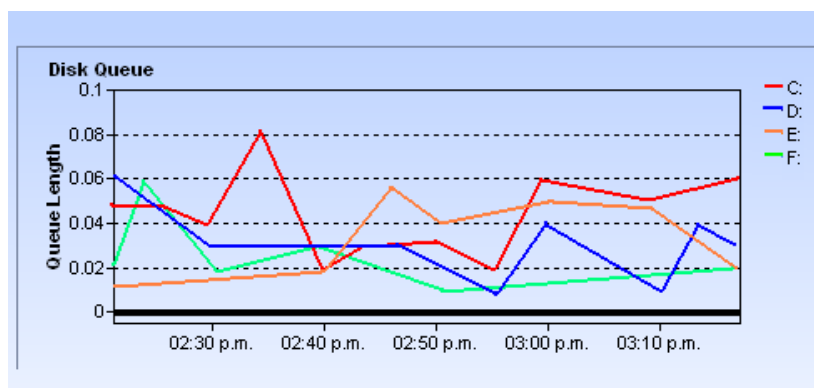
- Porcentaje de espacio disponible en el disco
- Longitud promedio de la cola de espera de utilización del disco

Al realizar el monitoreo de estos índices se puede determinar donde existe una alta carga de utilización de disco y que tanto afecta el rendimiento.

4.5.3.1 Longitud de espera por uso de los discos

Un indicador que ayuda en el análisis del estatus de rendimiento de los discos es la longitud de la cola de servicios en espera para ser atendidos, en la cual los valores aceptables no son más de 3 unidades por disco¹⁴ y en la figura siguiente podemos ver que los valores están por arriba de este parámetro, lo que sugiere una alta actividad del disco.

Figura 39. Longitud de peticiones de disco en cola de espera.



¹⁴ Algunos sistemas son capaces de tolerar hasta 5 unidades en espera de utilización del disco, antes de indicar una ineficiencia por solicitudes en espera a ser atendidos.

4.5.3.2 Disponibilidad de almacenamiento en el disco

La disponibilidad de almacenamiento en el disco, es de utilidad para determinar el área en la cual se puede hacer uso del disco y proyección de requerimientos futuros.

Generalmente debe de haber suficiente espacio para registrar las operaciones y datos, lo que lleva a realizar una proyección y tendencia de espacio disponible en el futuro, para contar siempre con un tamaño ideal de almacenamiento. ¿Qué tanto espacio debe de existir siempre en un disco?, es un tema un poco complicado, pues la información y área requerida depende de las necesidades propias de cada empresa, pero lo que sí se debe de tener en cuenta es cuanto espacio se utiliza diariamente y cuanto tiempo lleva realizar la compra o adquisición de más espacio para el sistema, hasta implementar el aumento del disco. Por lo cual es necesario medir constantemente el crecimiento de espacio utilizado y así proyectar las necesidades del disco.

Generalmente, los sistemas operativos cuentan con áreas temporales para cálculos especiales y procesamiento de archivos de manera temporal, las cuales con el transcurso de tiempo y forma de uso, pueden contener archivos temporales o de respaldo que no son necesarios que pueden ser eliminados con ayuda de programas específicos para buscar y eliminar estos tipos de archivos.

Es recomendable realizar por lo menos una vez a la semana la depuración de las áreas temporales, y sugerir a los usuarios que almacenan información de archivos para eliminar y depurar de la computadora a la red, para tener un copia de respaldo, para referencias futuras.

Figura 40. Espacio ocupado y disponible en el sistema de disco, representado en gigabytes



Como se puede observar, el porcentaje de espacios disponibles de cada disco es el siguiente

Tabla IX Detalle de porcentajes disponibles en discos

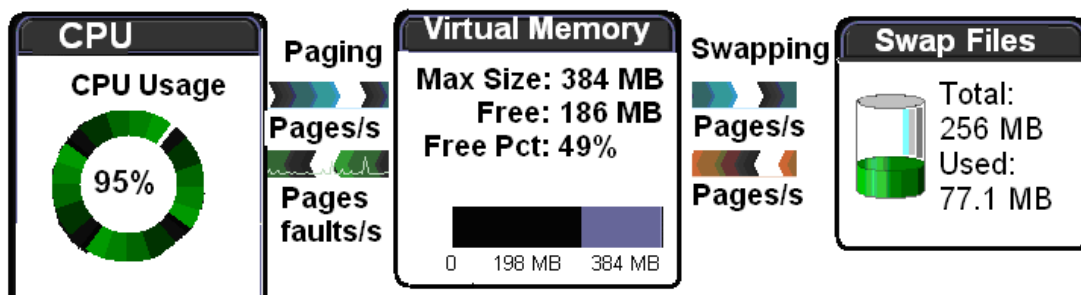
DISCO	C	D	F	G
TAMAÑO EN GB	4.00	9.75	5.75	14.25
TAMAÑO OCUPADO	3.00	7.25	4.25	10.00
PORCENTAJE DISPONIBLE	25.00	22.00	26.00	30.00

En resumen se tiene un 27.41% disponible del total de espacio (33.75GB).

4.5.3.3 Excesivo swap del disco

Cuando el sistema cuenta con un gran número de procesos, la actividad de paginación y swap puede dominar el rendimiento del sistema.

Figura 41. Actividad de lectura y escritura de disco por paginación y swap



Determinar cuando el sistema tiene una sobrecarga de paginación y swap requiere de un punto de partida. Es necesario tomar varias muestras en diferentes condiciones para determinar el punto de partida para una implementación específica. Por ejemplo, se puede reiniciar el sistema y ejecutar un chequeo de paginación y swap, con determinada cantidad de procesos en ejecución y realizarlo nuevamente durante un periodo de uso intenso del sistema, durante un periodo de alta actividad de tráfico de red y especialmente cuando la carga es alta y se está experimentando un bajo rendimiento.

Como se puede observar la cantidad de *swap* que realiza el sistema es de alrededor de 27 *swap* por segundo, hacia el disco (o sea escritura de datos) el cual es un factor alto, y el *swap* de lectura es de 6 *swap* por segundo. Esto da una idea que existe un retardo de *swap*, pero es principalmente porque la memoria no es suficiente.

4.6 Análisis de la red

La necesidad de estar comunicados para compartir información y solicitar servicios que otros sistemas proporcionan, etc., son aspectos indispensables y fundamentales para cumplir a cabalidad con las tareas diarias, ya que generalmente se cuenta con servidores que están conectados por medio de diferentes redes, y la ejecución de sus actividades son dependientes del rendimiento de la red.

El trabajo de esta sección está orientado específicamente cuando los problemas de rendimiento de red se refieren directamente al manejo y administración de la transmisión.

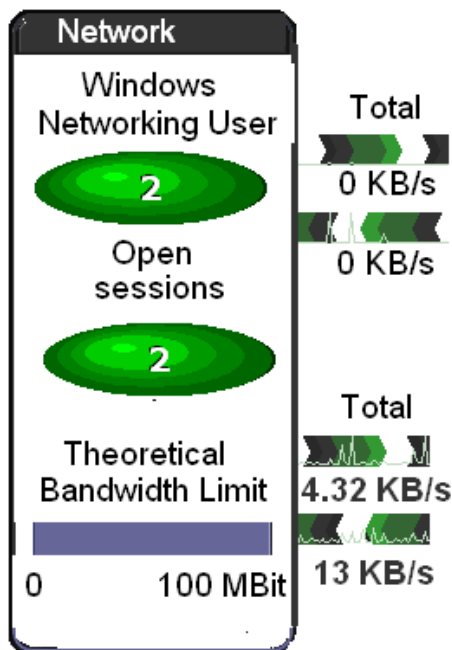
En un ambiente servidor, el objetivo de la red es enviar tanta información de datos como sea posible en un tiempo aceptable de respuesta sobre la red mientras consume la menor cantidad de recursos del sistema operativo (CPU, memoria, etc.). El métrico de mayor uso para evaluar el funcionamiento de la red es el rendimiento de trabajo, el cual usualmente es medido en megabits por segundo (Mbps) o gigabits por segundo (Gpps).

4.6.1 Consecuencias de un mal rendimiento de la red

El principal motivo de llevar a cabo un análisis de problemas de comunicación de red, es debido a que puede afectar el rendimiento del sistema operativo, ya sea por espera de un servicio de red, solicitud de datos, etc., requeridos a través de la red y que de manera errónea se concluye que el sistema operativo donde se está realizando las tareas es ineficiente, cuando en realidad el problema es externo al área de trabajo.

4.6.2 Aspectos a monitorear por el bajo rendimiento de la red

Figura 42. Estatus de descripciones de conexiones por red



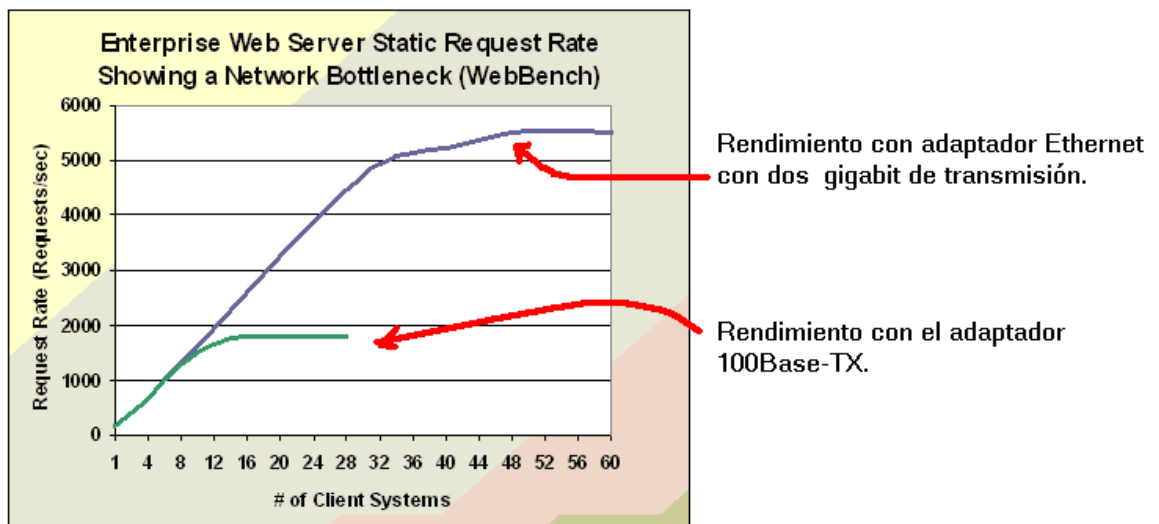
Cuando se está experimentando problemas de rendimiento y se tiene la sospecha de que el problema es a causa de la red, posterior a determinar que no son problemas de memoria, procesador o disco, es necesario realizar un análisis del estatus de conexión del servidor.

Estos datos son los primeros a evaluar, para entrar en detalle sobre determinado componente en el rendimiento de red.

Descripción de cantidad de conexiones de red, ancho de banda, conexiones *Ethernet* y estadísticas de flujo de transferencia de información.

La siguiente figura describe un cuello de botella generado al realizar un cambio de adaptador *Ethernet* de dos gigabit, por un adaptador de red 100Base-TX. Se muestra la reducción de la tasa de respuesta por segundos, al incrementarse la cantidad de usuarios conectados al sistema, en comparación del adaptador previo.

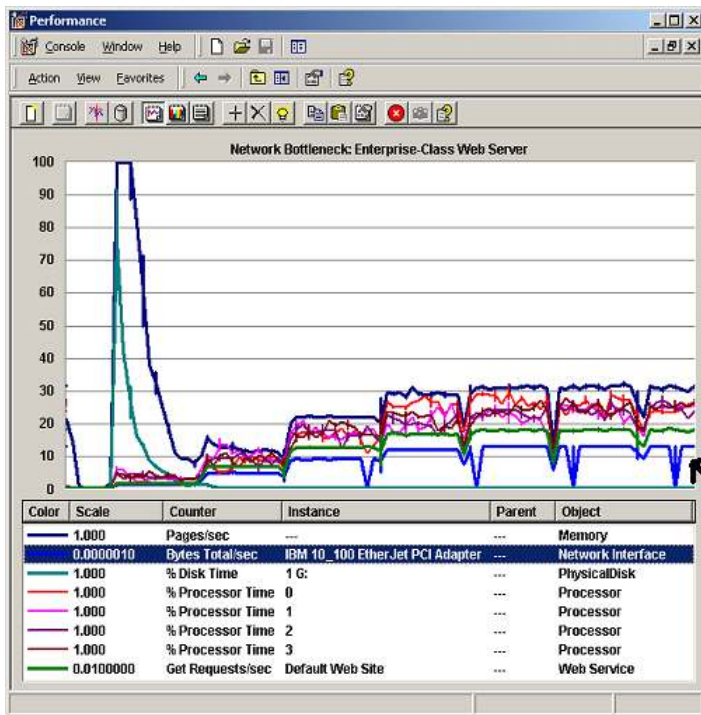
Figura 43. Monitoreo de un cuello de botella en la red



Cuello de botella en la red producido como consecuencia de un cambio de adaptador en el servidor

En la figura siguiente se muestra del rendimiento tomada durante la detección del cuello de botella, ocurrido durante el cambio del adaptador. El cuello de botella es creado por una limitante del ancho de banda disponible de transmisión de 100Mbps.

Figura 44. Cuello de botella del adaptador 100Base-TX, tomado con la herramienta *Performance Monitor* de Windows®

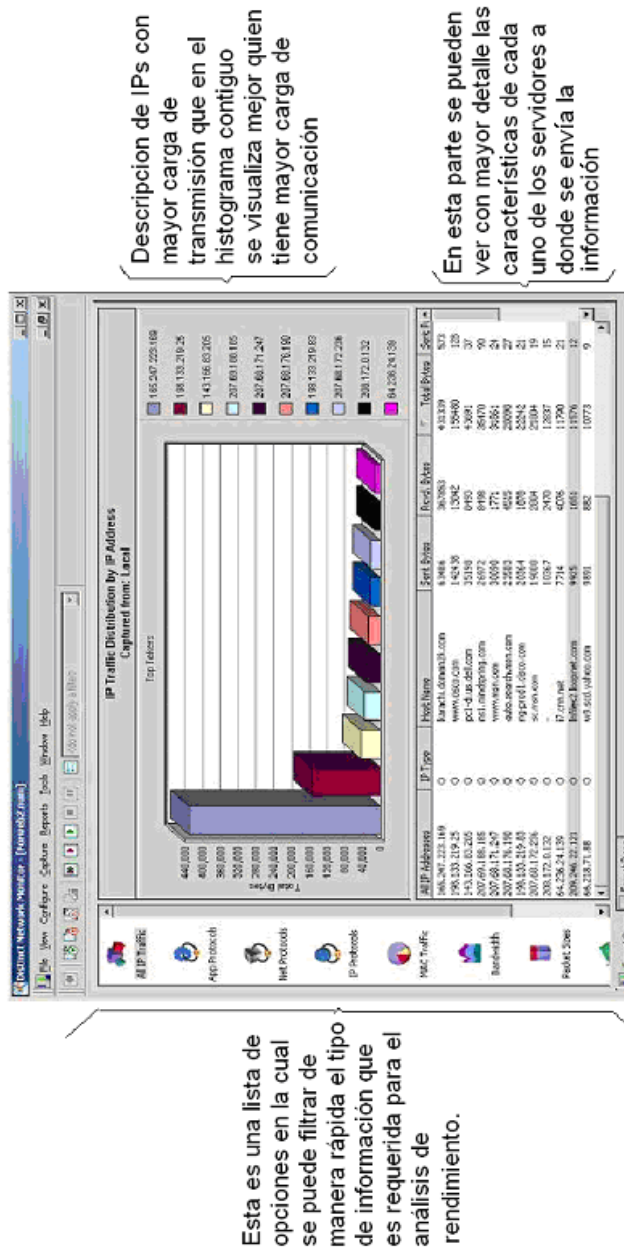


Rendimiento del adaptador en Bytes/segundos, el cual es en promedio de 10 Bytes/segundo

En este caso, es necesario instalar nuevamente la tarjeta *Ethernet* de dos gigabit, para corregir el problema debido a la cantidad de conexiones en el servidor.

Las herramientas de monitoreo en la parte de redes, proporcionan información de mucha utilidad, pues de manera sencilla es posible determinar a detalle, el comportamiento del tráfico que pasa a través del servidor que se está monitoreando.

Figura 45. Descripción de opciones que detalla la herramienta de monitoreo *Distinct Network Monitor*



La figura 45, muestra una opción de la herramienta de monitoreo de red llamada *Distinct Network Monitor*, Dentro de sus muchas funciones, permite visualizar de manera gráfica las estadísticas de los parámetros más importantes al momento de evaluar el sistema de red. Como en este caso el tráfico generado por la protocolo IP, y tendencias de cantidad de información a través de diferentes servidores. En la parte inferior, muestra detalles más específicos por tipo de servidor y al lado izquierdo, diferentes opciones de gráficas disponibles para la representación de la información.

En la figura 46, se puede observar como la información registrada y tabulada puede proporcionar de manera rápida y fácil, diferentes gráficas sobre el comportamiento general del sistema de red y la interrelación entre cada uno de los parámetros de mayor uso en la evaluación del rendimiento.

Con esta herramienta de monitoreo de red, se puede saber cuál es la cantidad máxima de unidades de transmisión, ordenando la información de bytes transmitidos. En este caso es de 12.6 Mbytes/segundo

CONCLUSIONES

1. Cada sistema operativo cuenta con variables de control y registro que son utilizados para medir el rendimiento y disponibilidad de los recursos, estas variables constantemente actualizan información por el sistema operativo sobre estadísticas de rendimiento y desempeño.
2. Los sistemas operativos para su evaluación, dividen sus recursos en objetos, y los parámetros utilizados para medir y evaluar el comportamiento de los objetos son llamados contadores que tienen límites para identificar los niveles aceptables sobre su desempeño llamados umbrales. Cuando se cuentan con más objetos de un mismo tipo, éstos son clasificados como instancias.
3. Los principales recursos de un sistema operativo que afectan directamente el rendimiento es la limitación de memoria, la sobrecarga del procesador, la velocidad de respuesta de los discos y el tráfico en la red.

4. La utilización de los comandos propios en conjunto con las herramientas de monitoreo permiten conocer la forma en que los sistemas operativos operan y las deficiencias en determinadas áreas que afectan el rendimiento, como los cambios necesarios para mejorar el rendimiento general del sistema.

5. Las herramientas de monitoreo facilitan la interpretación y comprensión del comportamiento de los sistemas operativos y detectan las áreas con problemas de manera más fácil, como también proporcionan la ayuda necesaria para mejorar las áreas en conflicto.

RECOMENDACIONES

1. Consultar variables, índices y parámetros de cada sistema operativo que miden el rendimiento en las guías de administración, previo a realizar el proceso de medición y afinamiento con el fin de identificar cuales son los límites y valores ideales y hasta que punto pueden ser modificados.
2. Para que puedan surgir efecto algunas modificaciones de los parámetros del sistema operativo es necesario reiniciar el sistema, por lo que se debe de programar y planificar el proceso de modificación y cambio a los variables del sistema, sin realizar más de un cambio al mismo tiempo, para revertirlo de ser necesario, por empeoramiento del mismo.
3. Observar continuamente el comportamiento del sistema operativo que presenta deficiencias en rendimiento, para identificar las áreas a mejorar y establecer políticas de ajustes y que mejor que la herramientas de monitoreo gráfico, para facilitar el análisis.
4. Establecer un orden de monitoreo, empezando por la memoria, procesador, disco y finalmente los componentes de redes involucrados en el tráfico de la información y comunicación, debido a su grado de importancia en el proceso del afinamiento.

5. Tomar en cuenta al momento de realizar el afinamiento de un sistema operativo, que no se puede expandir la capacidad más allá de la capacidad del *hardware* más lento en uso, por lo cual a veces es necesario adquirir más *hardware*, que sea más grande en capacidad, más rápido o de mejor tecnología.

BIBLIOGRAFÍA

1. Alonzo Lemus, Sergio Rodolfo, Consideraciones para el afinamiento y monitoreo de bases de datos relacionales. Tesis Ing. Ciencias y Sistemas. Universidad de San Carlos de Guatemala, Facultad de Ingeniería, 1999. 226pp.
2. *Appendix A. IRIX Kernel Tunable Parameters*
http://www.cepba.upc.es/docs/sgi_doc/SGI_Admin/books/IA_ConfiOps/sgi_html/apa.html
Agosto 2, 2003
3. *Common Performance Problems*
http://docs.sun.com/source/817-2180-10/pt_chap8.html
Mayo 7, 2004
4. Girón Miranda, José Daniel, Criterios y herramientas para la administración del tráfico en las redes de área local. Tesis Ing. Ciencias y Sistemas. Universidad de San Carlos de Guatemala, Facultad de Ingeniería, 2003. 101pp.
5. Guevara Castillo, Francisco Javier. Medición, evaluación y control del rendimiento en sistemas operativos. Tesis Ing. Ciencias y Sistemas. Universidad de San Carlos de Guatemala, Facultad de Ingeniería, 1990. 107pp.
6. Herramientas de monitoreo
<http://www2.ing.puc.cl/~msc/apuntes/cap5/cap5.htm>
Agosto 25, 2003

7. Jain, Raj. ***The art of computer systems performance analysis: techniques for experimental design measurements.*** Estados Unidos: John Wiley Y Sons. Inc, 1980 pp.230
8. Kant, Krishna. ***Introduction to computer system performance evaluation.*** Estados Unidos: McGraw, 1980. pp.230
9. *Microsoft Corporation* 1993-2003. **Biblioteca de consulta Encarta 2004.**
10. *Operating System Overview*
http://www.cs.arizona.edu/computer.help/policy/DIGITAL_unix/AA-Q0R3E-TET1_html/tune2.html
Julio 15, 2003
11. *Performance and Tuning of the UNIX Operating System*
<http://cs.felk.cvut.cz/~zemanekp/datasem/datasem.htm>
Junio 1, 2003
12. Puigjaner, Ramón y otros. **Evaluación y explotación de sistemas informáticos.** España: Síntesis, s.a. pp.254
13. *TCP Tuning Guide*
<http://www-didc.lbl.gov/TCP-tuning/>
Mayo 15, 2003
14. *WebLogic Server Performance and Tuning*
Tuning Hardware, Operating System, and Network Performance
<http://edocs.bea.com/wls/docs70/perform/HWTuning.html>
Mayo 7, 2003