

UNIVERSIDAD DE MÁLAGA
ESCUELA TÉCNICA SUPERIOR DE
INGENIERÍA INFORMÁTICA



PROYECTO FIN DE CARRERA:
**CARACTERIZACIÓN Y MODELADO POR TIPO DE
OBJETO DE TRÁFICO WEB**

INGENIERÍA INFORMÁTICA

Málaga, 2005

Mariano Rodríguez Tena

UNIVERSIDAD DE MÁLAGA
ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA DE INFORMÁTICA
INGENIERÍA TÉCNICA EN INFORMÁTICA DE GESTIÓN

Reunido el tribunal examinador en el día de la fecha, constituido por:

Presidente D/D^a _____

Secretario D/D^a _____

Vocal D/D^a _____

para juzgar el Proyecto Fin de Carrera titulado:

**CARACTERIZACIÓN Y MODELADO POR TIPO DE
OBJETO DE TRÁFICO WEB.**

del alumno D. Mariano Rodríguez Tena
dirigido por D. Francisco Javier González Cañete (Tecnología Electrónica)
 y D. Antonio Nebro Urbaneja (Lenguajes y Ciencias de la Computación)

ACORDÓ POR _____ OTORGAR LA CALIFICACIÓN DE _____

y para que conste, se extiende firmada por los comparecientes del Tribunal, la presente diligencia.

Málaga, a ____ de _____ de 2005

El Presidente:

El Secretario:

El Vocal:

Fdo. _____

Fdo. _____

Fdo. _____

ÍNDICE

ÍNDICE	I
LISTADO DE FIGURAS	III
LISTADO DE TABLAS	IV
CAPITULO 1: INTRODUCCIÓN	1
1.1. DESCRIPCIÓN DE LAS TRAZAS	3
CAPITULO 2: HERRAMIENTAS DEL PROYECTO	4
2.1. PROGRAMA DE TROCEADO	4
2.1.1. <i>Los archivos de trazas</i>	4
2.1.2. <i>Particionado y clasificación</i>	7
2.1.3. <i>Ejemplo básico</i>	9
2.1.4. <i>Formas de ejecución</i>	10
2.2. FRECUENCIAS URL	12
2.2.1. <i>Ejemplo básico</i>	12
2.2.2. <i>Formas de ejecución</i>	14
2.3. ASIGNADOR DE CODIGOS HASH A URL	16
2.3.1. <i>Ejemplo básico</i>	16
2.3.2. <i>Formas de ejecución</i>	18
2.4. SIMULADOR DE PROXY-CACHE	20
2.4.1. <i>Políticas de reemplazo</i>	20
2.4.2. <i>Simulación simple y por lotes</i>	21
2.4.3. <i>Archivo de propiedades</i>	21
2.4.3.1. <i>Ajuste del tipo de cache</i>	22
2.4.3.2. <i>Ajuste del tamaño de cache</i>	23
2.4.3.3. <i>Ajuste de la ubicación de los ficheros</i>	23
2.4.3.4. <i>Ajuste de los tipos de objetos a almacenar</i>	24
2.4.4. <i>Ejemplo básico</i>	24
2.4.5. <i>Formas de ejecución</i>	25
2.4.6. <i>Javadoc de la aplicación</i>	26
2.5. PROGRAMA PRINCIPAL	27
2.5.1. <i>Inicio</i>	27
2.5.2. <i>Opciones del menu principal</i>	28
2.5.2.1. <i>Cambiar el directorio de trabajo</i>	29
2.5.2.2. <i>Porcentaje de frecuencias</i>	30
2.5.2.3. <i>Porcentaje de bytes</i>	31
2.5.2.4. <i>Media, mediana, y desviación típica</i>	31
2.5.2.5. <i>Histograma de los servidores más usados</i>	32
2.5.2.6. <i>Gráfica de las URL más visitadas</i>	33
2.5.2.7. <i>Estudiar una traza en general</i>	34
2.5.2.8. <i>Distancia entre peticiones a un objeto</i>	34
2.5.2.9. <i>Tamaño entre peticiones a un objeto</i>	34
2.5.2.10. <i>Tiempo entre peticiones a un objeto</i>	35
2.5.2.11. <i>Obtención del campo de una traza</i>	36
2.5.2.12. <i>Función de distribución de los tamaños</i>	37
2.5.2.13. <i>Simular una proxy-cache</i>	38
2.5.2.14. <i>Salir</i>	38
2.5.3. <i>Ejecución por lotes</i>	38
CAPÍTULO 3: ESTUDIO ESTADÍSTICO	39

3.1. ESTUDIO DE TRAZAS EN GENERAL	39
3.1.1. Servidor RTP	39
3.1.1.1. Resumen	42
3.1.2. Servidor PA	44
3.1.2.1. Resumen	46
3.2. ESTUDIO DE TRAZAS SEGÚN CLASES Y TIPOS DE OBJETOS	48
3.2.1. Estudio por clase de objeto	48
3.2.1.1. Servidor RTP	48
3.2.1.2. Servidor PA	77
3.2.1.3. Resumen	104
3.2.2. Estudio por tipo de objeto	107
3.2.2.1. Servidor RTP	108
3.2.2.2. Servidor PA	111
3.2.2.3. Resumen	113
CAPITULO 4: CARACTERIAZACIÓN DEL TRÁFICO	116
4.1. ESTUDIO DE SUCESIVAS PETICIONES	116
4.1.1. Distancias entre peticiones	116
4.1.2. Tamaños entre peticiones	124
4.1.3. Tiempos entre peticiones	131
4.1.4. Conclusiones	139
4.2. CURVAS DE DISTRIBUCIÓN DE LOS TAMAÑOS	140
4.2.1. Servidor RTP	140
4.2.2. Servidor PA	142
4.2.3. Conclusiones	143
4.3. FRECUENCIA DE PETICIONES	145
4.3.1. Conclusiones	147
CAPÍTULO 5: SIMULACIONES PROXY-CACHES	149
5.1. POLÍTICAS DE REEMPLAZO	150
5.1.1. Política de reemplazo LRU (Least Recently Used)	150
5.1.2. Política de reemplazo LFU (Least Frequently Used)	150
5.1.3. Política de reemplazo híbrida	151
5.2. RESULTADOS	151
5.3. CONCLUSIONES	163
CAPÍTULO 6: CONCLUSIONES Y LÍNEAS FUTURAS	165
APÉNDICE A: CLASIFICACIÓN DE EXTENSIONES	167
BIBLIOGRAFÍA	169

LISTADO DE FIGURAS

FIGURA 2.1. EJEMPLO DE USO DE LA APLICACIÓN DE TROCEADO.	10
FIGURA 2.2. EJEMPLO DE EJECUCIÓN DE LA APLICACIÓN DE FREURL.EXE.	14
FIGURA 2.3. EJEMPLO DE EJECUCIÓN DE LA APLICACIÓN DISTURL2.EXE.	18
FIGURA 2.4. ELECCIÓN DEL DIRECTORIO DE TRABAJO EN MATLAB.	28
FIGURA 2.5. MENÚ PRINCIPAL DE LA APLICACIÓN.	29
FIGURA 2.6. ELECCIÓN DEL DIRECTORIO DE TRABAJO.	30
FIGURA 2.7. VENTANA MODAL PARA LA ELECCIÓN DE LOS TIPOS DE OBJETOS A PROCESAR.	31
FIGURA 2.8. VENTANA MODAL PARA LA ELECCIÓN DEL RANGO DE CÓDIGOS HTTP.	32
FIGURA 2.9. SELECCIÓN DEL NÚMERO DE TRAZAS A PROCESAR.	33
FIGURA 2.10. ELECCIÓN EL NUMERO HASH DE OBJETO Y FICHERO A ESTUDIAR.	35
FIGURA 2.11. ELECCIÓN DEL TIPO DE GRAFICA A MOSTRAR.	36
FIGURA 2.12. ELECCIÓN DE LOS CAMPOS DE TRAZA A OBTENER.	37
FIGURA 3.1. GRÁFICO DE FRECUENCIAS DE LOS CÓDIGOS HTTP.	43
FIGURA 3.2. GRÁFICO DE FRECUENCIAS DE LOS MÉTODOS HTTP.	43
FIGURA 3.3. GRÁFICO DE FRECUENCIAS DE LOS CÓDIGOS HTTP.	47
FIGURA 3.4. GRÁFICO DE FRECUENCIAS DE LOS MÉTODOS HTTP.	47
FIGURA 3.5. PORCENTAJE DE BYTES DE LAS CLASES DE OBJETOS.	105
FIGURA 3.6. TAMAÑO MEDIO DE LAS CLASES DE OBJETOS.	106
FIGURA 3.7. TIEMPOS DE DESCARGA DE LAS CLASES DE OBJETO.	106
FIGURA 3.8. HISTOGRAMA DE COMPARACIÓN DE LAS TASAS DE TRANSFERENCIA.	107
FIGURA 3.9. FRECUENCIAS DE PETICIONES DE LOS PRINCIPALES TIPOS DE OBJETOS.	114
FIGURA 3.10. PORCENTAJE DE BYTES DE LOS PRINCIPALES TIPOS DE OBJETO.	115
FIGURA 3.11. TAMAÑO MEDIO DE LOS PRINCIPALES TIPOS DE OBJETO.	115
FIGURA 4.1. DISTANCIAS ENTRE PETICIONES A 1 PARA EL DÍA 22/03/04.	117
FIGURA 4.2. DISTANCIAS ENTRE PETICIONES A 1 PARA EL DÍA 23/03/04.	117
FIGURA 4.3. DISTANCIAS ENTRE PETICIONES A 1 PARA EL DÍA 24/03/04.	118
FIGURA 4.4. DISTANCIAS ENTRE PETICIONES A 2 PARA EL DÍA 22/03/04.	118
FIGURA 4.5. DISTANCIAS ENTRE PETICIONES A 2 PARA EL DÍA 23/03/04.	119
FIGURA 4.6. DISTANCIAS ENTRE PETICIONES A 2 PARA EL DÍA 24/03/04.	119
FIGURA 4.7. DISTANCIAS ENTRE PETICIONES A 3 PARA EL DÍA 22/03/04.	120
FIGURA 4.8. DISTANCIAS ENTRE PETICIONES A 3 PARA EL DÍA 23/03/04.	120
FIGURA 4.9. DISTANCIAS ENTRE PETICIONES A 3 PARA EL DÍA 24/03/04.	121
FIGURA 4.10. DISTANCIAS ENTRE PETICIONES A 3 PARA EL DÍA 25/03/04.	121
FIGURA 4.11. DISTANCIAS ENTRE PETICIONES A 3 PARA EL DÍA 26/03/04.	122
FIGURA 4.12. DISTANCIAS ENTRE PETICIONES A 3 PARA EL DÍA 27/03/04.	122
FIGURA 4.13. DISTANCIAS ENTRE PETICIONES A 3 PARA EL DÍA 28/03/04.	123
FIGURA 4.14. DISTANCIAS ENTRE PETICIONES A 4 PARA EL DÍA 22/03/04.	123
FIGURA 4.15. TAMAÑOS ENTRE PETICIONES A 1 PARA EL DÍA 22/03/04.	124
FIGURA 4.16. TAMAÑOS ENTRE PETICIONES A 1 PARA EL DÍA 23/03/04.	125
FIGURA 4.17. TAMAÑOS ENTRE PETICIONES A 1 PARA EL DÍA 24/03/04.	125
FIGURA 4.18. TAMAÑOS ENTRE PETICIONES A 2 PARA EL DÍA 22/03/04.	126
FIGURA 4.19. TAMAÑOS ENTRE PETICIONES A 2 PARA EL DÍA 23/03/04.	126
FIGURA 4.20. TAMAÑOS ENTRE PETICIONES A 2 PARA EL DÍA 24/03/04.	127
FIGURA 4.21. TAMAÑOS ENTRE PETICIONES A 3 PARA EL DÍA 22/03/04.	127
FIGURA 4.22. TAMAÑOS ENTRE PETICIONES A 3 PARA EL DÍA 23/03/04.	128
FIGURA 4.23. TAMAÑOS ENTRE PETICIONES A 3 PARA EL DÍA 24/03/04.	128
FIGURA 4.24. TAMAÑOS ENTRE PETICIONES A 3 PARA EL DÍA 25/03/04.	129
FIGURA 4.25. TAMAÑOS ENTRE PETICIONES A 3 PARA EL DÍA 26/03/04.	129
FIGURA 4.26. TAMAÑOS ENTRE PETICIONES A 3 PARA EL DÍA 27/03/04.	130
FIGURA 4.27. TAMAÑOS ENTRE PETICIONES A 3 PARA EL DÍA 28/03/04.	130
FIGURA 4.28. TAMAÑOS ENTRE PETICIONES A 4 PARA EL DÍA 22/03/04.	131
FIGURA 4.29. TIEMPOS ENTRE PETICIONES A 1 PARA EL DÍA 22/03/04.	132
FIGURA 4.30. TIEMPOS ENTRE PETICIONES A 1 PARA EL DÍA 23/03/04.	132

FIGURA 4.31. TIEMPOS ENTRE PETICIONES A 1 PARA EL DÍA 24/03/04.....	133
FIGURA 4.32. TIEMPOS ENTRE PETICIONES A 2 PARA EL DÍA 22/03/04.....	133
FIGURA 4.33. TIEMPOS ENTRE PETICIONES A 2 PARA EL DÍA 23/03/04.....	134
FIGURA 4.34. TIEMPOS ENTRE PETICIONES A 2 PARA EL DÍA 24/03/04.....	134
FIGURA 4.35. TIEMPOS ENTRE PETICIONES A 3 PARA EL DÍA 22/03/04.....	135
FIGURA 4.36. TIEMPOS ENTRE PETICIONES A 3 PARA EL DÍA 23/03/04.....	135
FIGURA 4.37. TIEMPOS ENTRE PETICIONES A 3 PARA EL DÍA 24/03/04.....	136
FIGURA 4.38. TIEMPOS ENTRE PETICIONES A 3 PARA EL DÍA 25/03/04.....	136
FIGURA 4.39. TIEMPOS ENTRE PETICIONES A 3 PARA EL DÍA 26/03/04.....	137
FIGURA 4.40. TIEMPOS ENTRE PETICIONES A 3 PARA EL DÍA 27/03/04.....	137
FIGURA 4.41. TIEMPOS ENTRE PETICIONES A 3 PARA EL DÍA 28/03/04.....	138
FIGURA 4.42. TIEMPOS ENTRE PETICIONES A 4 PARA EL DÍA 22/03/04.....	138
FIGURA 4.43. CURVAS DE DISTRIBUCIÓN DE LOS TAMAÑOS PARA EL SERVIDOR RTP	141
FIGURA 4.44. CURVAS DE DISTRIBUCIÓN DE LOS TAMAÑOS PARA EL SERVIDOR PA	143
FIGURA 4.45. EJEMPLO DE GRÁFICAS DE UNA DISTRIBUCIÓN ZIPF	145
FIGURA 4.46. FRECUENCIA DE PETICIONES PARA EL 22/03/04 Y 23/03/04.....	146
FIGURA 4.47. FRECUENCIA DE PETICIONES PARA EL 23/03/04 Y 24/03/04.....	146
FIGURA 4.48. FRECUENCIA DE PETICIONES PARA EL 26/03/04 Y 27/03/04.....	147
FIGURA 4.49. FRECUENCIA DE PETICIONES PARA EL 28/03/04.....	147
FIGURA 5.1. RENDIMIENTO DE LA PROXY-CACHE PARA 22/03/04	153
FIGURA 5.2. RENDIMIENTO DE LA PROXY-CACHE PARA 23/03/04	155
FIGURA 5.3. RENDIMIENTO DE LA PROXY-CACHE PARA 24/03/04	156
FIGURA 5.4. RENDIMIENTO DE LA PROXY-CACHE PARA 25/03/04	158
FIGURA 5.5. RENDIMIENTO DE LA PROXY-CACHE PARA 26/03/04	160
FIGURA 5.6. RENDIMIENTO DE LA PROXY-CACHE PARA 27/03/04	161
FIGURA 5.7. RENDIMIENTO DE LA PROXY-CACHE PARA 28/03/04	163

LISTADO DE TABLAS

TABLA 2.1. ASOCIACIÓN DE VALORES A LAS ETIQUETAS DE TRATAMIENTO DEL DATO.....	7
TABLA 2.2. ASOCIACIÓN DE VALORES A LAS ETIQUETAS DE MÉTODO DE LA PETICIÓN.	8
TABLA 2.3. ASOCIACIÓN DE VALORES PARA LAS DISTINTAS ETIQUETAS DE JERARQUÍAS DE DATOS.	8
TABLA 2.4. CLASIFICACIÓN DE EXTENSIONES RELAZADA POR LA APLICACIÓN DE PRE-PROCESAMIENTO.	9
TABLA 2.5. LISTADO DE PROPIEDADES QUE DEBEN SER CONFIGURADAS.....	22
TABLA 3.1. ESTUDIO GENERAL PARA LAS TRAZAS DEL 22/03/04 DEL SERVIDOR RTP.....	40
TABLA 3.2. ESTUDIO GENERAL PARA LAS TRAZAS DEL 23/03/04 DEL SERVIDOR RTP.....	40
TABLA 3.3. ESTUDIO GENERAL PARA LAS TRAZAS DEL 24/03/04 DEL SERVIDOR RTP.....	41
TABLA 3.4. ESTUDIO GENERAL PARA LAS TRAZAS DEL 25/03/04 DEL SERVIDOR RTP.....	41
TABLA 3.5. ESTUDIO GENERAL PARA LAS TRAZAS DEL 26/03/04 DEL SERVIDOR RTP.....	41
TABLA 3.6. ESTUDIO GENERAL PARA LAS TRAZAS DEL 27/03/04 DEL SERVIDOR RTP.....	42
TABLA 3.7. ESTUDIO GENERAL PARA LAS TRAZAS DEL 28/03/04 DEL SERVIDOR RTP.....	42
TABLA 3.8. ESTADÍSTICOS DE RESUMEN PARA TODAS LAS TRAZAS DEL SERVIDOR RTP	43
TABLA 3.9. ESTUDIO GENERAL PARA LAS TRAZAS DEL 22/03/04 DEL SERVIDOR PA.....	44
TABLA 3.10. ESTUDIO GENERAL PARA LAS TRAZAS DEL 23/03/04 DEL SERVIDOR PA.....	44
TABLA 3.11. ESTUDIO GENERAL PARA LAS TRAZAS DEL 24/03/04 DEL SERVIDOR PA.....	45
TABLA 3.12. ESTUDIO GENERAL PARA LAS TRAZAS DEL 25/03/04 DEL SERVIDOR PA.....	45
TABLA 3.13. ESTUDIO GENERAL PARA LAS TRAZAS DEL 26/03/04 DEL SERVIDOR PA.....	45
TABLA 3.14. ESTUDIO GENERAL PARA LAS TRAZAS DEL 27/03/04 DEL SERVIDOR PA.....	46
TABLA 3.15. ESTUDIO GENERAL PARA LAS TRAZAS DEL 28/03/04 DEL SERVIDOR PA.....	46
TABLA 3.16. ESTADÍSTICOS DE RESUMEN PARA EL SERVIDOR PA.....	47
TABLA 3.17: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 22/03/04	52

TABLA 3.18: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 23/03/04	56
TABLA 3.19: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 24/03/04	60
TABLA 3.20: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 25/03/04	64
TABLA 3.21: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 26/03/04	68
TABLA 3.22: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 27/03/04	72
TABLA 3.23: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 28/03/04	76
TABLA 3.24: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 22/03/04	80
TABLA 3.25: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 23/03/04	84
TABLA 3.26: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 24/03/04	88
TABLA 3.27: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 25/03/04	92
TABLA 3.28: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 26/03/04	96
TABLA 3.29: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 27/03/04	100
TABLA 3.30: ESTUDIO POR CLASE DE OBJETO DE LAS TRAZAS DEL 28/03/04	104
TABLA 3.31: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 22/03/04.....	108
TABLA 3.32: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 23/03/04.....	109
TABLA 3.33: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 24/03/04.....	109
TABLA 3.34: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 25/03/04.....	109
TABLA 3.35: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 26/03/04.....	110
TABLA 3.36: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 27/03/04.....	110
TABLA 3.37: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 28/03/04.....	110
TABLA 3.38: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 22/03/04.....	111
TABLA 3.39: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 23/03/04.....	111
TABLA 3.40: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 24/03/04.....	112
TABLA 3.41: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 25/03/04.....	112
TABLA 3.42: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 26/03/04.....	112
TABLA 3.43: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 27/03/04.....	113
TABLA 3.44: ESTUDIO DE LOS PRINCIPALES TIPOS DE OBJETO PARA LAS TRAZAS DEL 28/03/04.....	113
TABLA 4.1. COMPARACIÓN DE LA DISTRIBUCIÓN ACUMULADA DE LOS TAMAÑOS	144
TABLA 5.1. DIFERENCIAS DE EFICIENCIA ENTRE PROXY CACHES LFU Y LRU	164

CAPITULO 1: Introducción

La World Wide Web esta basada en el modelo cliente-servidor. Un cliente accede a los documentos en la Web a través de un navegador Web. El navegador envía una petición al servidor, el cual responde con el documento solicitado. Aunque la información puede ser almacenada en cualquier localización del mundo, el usuario tiene la ilusión de que esos datos están almacenados localmente.

Un servidor Web puede responder a peticiones desde múltiples clientes Web. La comunicación esta siempre en la forma petición-respuesta, y es siempre iniciada por el cliente. Clientes y servidores Web se comunican usando el Protocolo de Transferencia de Hipertexto (HTTP). HTTP esta montado sobre TCP, un protocolo orientado al flujo bidireccional de bytes de la capa de transporte.

Un usuario puede acceder a la información sobre la Web usando un navegador Web, tal como netscape o internet explorer. Cuando el usuario selecciona un documento (normalmente haciendo ‘click’ sobre un hiperlink), el navegador Web crea una petición que será enviada al correspondiente servidor Web. La petición incluye: el nombre del documento solicitado, expresado como una URL (Uniform Resource Locator); un conjunto de cabeceras de la petición de hipertexto, donde se incluye, por ejemplo, el formato de los datos que el cliente acepta; e información de la autenticidad del usuario, la cual dice al servidor si el cliente tiene permisos para recibir ese documento. Una vez la petición ha sido enviada al servidor Web, la maquina del cliente espera la respuesta. Cuando llega, el navegador la analiza. Dependiendo de esta, la maquina del cliente puede hacer otra respuesta al servidor, o mostrar el documento al usuario.

Por otro lado, un servidor Web opera como sigue: el servidor ‘escucha’ sobre un puerto (normalmente el puerto 80) en busca de una petición de un cliente Web para establecer una conexión TCP. Una vez la conexión TCP se ha abierto y el cliente ha hecho su petición, el servidor debe responder a esta. La respuesta incluye un código de estado que informa al cliente si su petición tuvo éxito. Si es así, la respuesta incluye el documento solicitado, si no es así, la razón del fallo es devuelto al cliente. Una vez el servidor Web ha enviado su respuesta y la conexión TCP se cierra con el cliente, el servidor repite el ciclo y comienza a ‘escuchar’ su próxima petición.

El propósito, por tanto, del servidor Web es proveer documentos a los clientes Web que los solicitan. Cada página Web puede consistir de múltiples documentos. Hasta la versión 1.0 de http, cada documento era solicitado de forma independiente al servidor.

Los servidores Web pueden ser configurados para registrar la información relativa a las peticiones de los clientes. Estos registros incluyen información tan diversa como el instante de tiempo en el que se solicita el documento, que tamaño tiene o de que tipo es, y pueden ser utilizados para que, a través de su análisis, pueda mejorarse el rendimiento de la Web.

La popularidad de la Web se ha incrementado dramáticamente hace unos pocos años. Hace poco más de una década el tráfico Web era inexistente. Hoy, el tráfico Web es uno de los componentes dominantes en Internet.

Hay muchas razones detrás de este crecimiento explosivo. Estas razones incluyen el fácil uso de la Web; la disponibilidad de interfaces gráficas de usuario para navegar; la existencia de editores y herramientas de soporte para la creación y publicación de documentos Web; una emergente tendencia de investigadores, instituciones educativas, y organizaciones comerciales para usar la Web como diseminador de información; y un continuo incremento en el número de hosts y usuarios en Internet.

Este crecimiento exponencial ha llevado a las demandas de infraestructura para difundir y acceder a documentos Web. Las nuevas demandas que han surgido debido a ese crecimiento han fomentado la búsqueda de soluciones a problemas tales como: grandes congestiones en la red, un bajo ancho de banda, altas latencias al momento de extraer un documento, sobrecarga en servidores, accesos simultáneos masivos a un servidor con temas atractivos que producen partición de la red, entre otros. Esto ha orientado la actividad de los investigadores a intentar mejorar el rendimiento de la Web y su escalabilidad. Para esto la clave a considerar es como reducir el volumen de tráfico producido por clientes y servidores Web, y como mejorar los tiempos de respuesta para los usuarios de WWW.

En este sentido se ha recurrido al uso de *caching* aplicado en la Web. El *caching* en la Web reduce las cargas en la red, en los servidores, y las latencias en el momento de extraer un documento remoto. Sin embargo, los sistemas de *caching* presentan varias problemáticas y

cuestiones de diseño que no son fáciles de resolver y que deben tenerse en cuenta para no producir resultados contradictorios al momento de utilizar cache en la Web.

Se hace fundamental entonces una sólida comprensión de la naturaleza y estructura del tráfico Web. Caracterizar y modelar el mismo es el objetivo principal de este proyecto, así como el desarrollo de las herramientas necesarias para posibilitar dicho estudio.

La memoria del proyecto se desarrolla en dos partes:

- i. La primera de ellas se enfoca al manual de usuario de las herramientas desarrolladas en este proyecto para la realización de los cálculos, teniendo como principal objetivo la total comprensión, por parte del lector, de la funcionalidad y utilización de todo el software. Explicaremos a lo largo del capítulo segundo las herramientas usadas para el pre-procesamiento de las trazas, la obtención de resultados estadísticos, y las simulaciones de proxy-caches.
- ii. La segunda parte se enfocará a mostrar y comentar los resultados obtenidos haciendo uso de todas las herramientas expuestas en la primera parte. Análisis estadístico de las clases y principales tipos de objetos formará el capítulo tercero. Los estudios de los espacios-tiempos-tamaños entre peticiones, las curvas de distribución de los tamaños de los objetos, y la aproximación de las frecuencias de peticiones de objetos en la Web a una distribución zipf formarán el capítulo cuarto. Por otro lado, el quinto se dedicará al estudio del rendimiento de las cache de los proxy con distintas políticas de reemplazo.

1.1. DESCRIPCIÓN DE LAS TRAZAS

En el estudio de este proyecto se han utilizado archivos de trazas obtenidos en <http://www.ircache.net>. Los archivos de trazas que se han tratado en concreto son los pertenecientes a los servidores proxy con el mayor y el menor flujo de información de entre todos los suministrados. Estos servidores proxy se encuentran en Research Triangle Park, Carolina del norte (y al que nos referiremos con las siglas RTP), y Palo Alto, California (al que nos referiremos con las siglas PA). El estudio se centra en las trazas que van del 22 de Marzo de 2004 al 28 de Marzo de 2004, y suman un total de 14.397.547 peticiones.

CAPITULO 2: Herramientas del proyecto

En este capítulo describiremos la funcionalidad y modo de uso de cada uno de los programas desarrollados en la elaboración de este proyecto.

Para la mejor comprensión de la funcionalidad del producto final, se irán exponiendo de forma ordenada cada una de estas herramientas. Primero con las que, si bien tienen una funcionalidad simple, sirven de base a otras con un fin inmediato. Seguidamente expondremos estas más elaboradas. Y finalizaremos con la aplicación principal cuya finalidad será la de coordinar el funcionamiento de todas las anteriores.

2.1. PROGRAMA DE TROCEADO

Esta aplicación implementada en C, sirve de base a todo el proyecto. Su necesidad radica en preparar y dar formato a los datos que mas tarde se tomarán como entrada por otras aplicaciones del proyecto. Estos datos de entrada aun sin preprocesar son los archivos de trazas. A lo largo de este apartado se explica cómo funciona esta aplicación, qué salida genera, y la forma de utilizarlo.

2.1.1. LOS ARCHIVOS DE TRAZAS

La Web esta basada en el modelo cliente-servidor. La comunicación esta siempre en la forma de par petición-respuesta y siempre iniciada por el cliente. Un cliente accede a documentos de la Web usando un navegador Web. Cuando el usuario selecciona un documento para recibir, el navegador crea una petición para ser enviada al correspondiente servidor Web. Cada página Web puede consistir en múltiples documentos (archivos). Actualmente, cada archivo se solicita por separado. Un servidor Web responde a cada petición que recibe desde los clientes Web. La respuesta del servidor incluye un código estado para informar al cliente del éxito (o no) de la petición. Si la petición tuvo éxito, entonces la respuesta incluye el documento solicitado. Si no tuvo éxito, una razón del fallo se retorna al cliente.

Los servidores Web pueden ser configurados para almacenar información acerca de todas las peticiones y respuestas procesadas por el servidor. Toda esta información forma los archivos de trazas. Estos son archivos de texto en los que cada línea contiene información sobre la

petición de un documento. Esta información se encuentra separada por un espacio en blanco (aunque podría ser cualquier otro carácter siempre y cuando sea el patrón general). En adelante nos referiremos a ella como ‘campo de una traza’. Los campos que encontramos en una traza son los siguientes:

1. *Timestamp*: es el instante de tiempo en que el socket del cliente se cierra.
2. *Elapsed time*: es el tiempo de la petición en milisegundos. Dicho de otra manera, es el tiempo comprendido entre las llamadas a las funciones `accept()` y `close()` del socket del cliente. Para el caso de conexiones HTTP persistentes, este es el tiempo entre la lectura del primer byte y la escritura del último.
3. Dirección del Cliente: Una dirección IP aleatoria que identificará al cliente. La correspondencia cliente-dirección permanece la misma para todas las peticiones de un fichero de trazas.
4. Etiqueta de tratamiento y código HTTP. La etiqueta de tratamiento describe como la petición fue tratada localmente. Estos son los posibles valores:
 - a. TCP_HIT: el objeto solicitado ya estaba en cache.
 - b. TCP_MEM_HIT: una copia del objeto se encontró en la memoria cache ‘Hot object’.
 - c. TCP_MISS: el objeto solicitado no se encontraba en cache.
 - d. TCP_REFRESH_HIT: el objeto estaba en cache, pero estaba ‘caducado’. Se realiza una solicitud *if-modified-since* y la respuesta es un 304-No modificado.
 - e. TCP_REF_FAIL_HIT: el objeto estaba en cache, pero estaba ‘caducado’. La petición para validar el objeto fallo y por tanto el objeto ‘caducado’ fue devuelto.
 - f. TCP_REFRESH_MISS: objeto estaba en cache, pero estaba ‘caducado’. Se realiza una solicitud *if-modified-since* y la respuesta es el nuevo objeto.
 - g. TCP_CLIENT_REFRESH_MISS: El cliente realizo una petición con la orden ‘no-cache’
 - h. TCP_IMS_HIT: El cliente realizo una solicitud del tipo *if-modified-since* y el resultado es que el objeto ya esta actualizado.
 - i. TCP_IMS_MISS: el cliente realiza una solicitud del tipo *if-modified-since* para un objeto ‘caducado’.

- j. TCP_SWAPFAIL_MISS: Se creyó que el objeto estaría en cache, pero no pudo ser accedido.
 - k. TCP_DENIED: el acceso fue denegado para esa solicitud.
5. Tamaño: Número de bytes escritos en el cliente.
 6. Método de petición: Es el método HTTP de la petición (GET, POST, o HEAD).
 7. URL: Es la URL (Uniform Resource) solicitada. Los argumentos CGI están seguidos por ‘?’.
 8. Identidad del usuario: Siempre ‘-’.
 9. Etiqueta de jerarquía de datos y hostname: Una descripción de como y de donde el objeto solicitado fue traído. La etiqueta de jerarquía de datos puede tener los siguientes valores:
 - a. NONE: Para TCP HIT, TCP fallos, peticiones *cachemgr* y todas las peticiones UDP, no hay jerarquía de información.
 - b. DIRECT: El objeto fue solicitado desde el servidor original.
 - c. FIRST_PARENT_MISS: The object was fetched from the parent with the fastest (possibly weighted) round trip time. El objeto fue traído desde el padre con la más rápida (posiblemente pesada) rueda de tiempo.
 - d. SIBLING_HIT: El objeto se trae desde una cache hermana la cual contesto con un TCP_HIT
 - e. PARENT_HIT: El objeto se solicita desde una cache padre y responde con un TCP_HIT
 - f. SINGLE_PARENT: El objeto se pidió desde el único padre apropiado para la URL dada.
 - g. CLOSEST_PARENT_MISS: Este padre fue elegido, porque incluía la más baja medida RTT para el servidor original.
 - h. CLOSEST_DIRECT: Las medidas RTT contestaron antes que cualquiera de los padres.
 - i. CACHE_DIGEST_HIT: El compañero fue elegido porque la *cache digest* predijo un *hit*. Esta opción fue más tarde reemplazada con el objetivo de distinguir entre padres y hermanos.
 - j. CLOSEST_PARENT: La opción del padre estuvo basada en las medidas RTT.

Además cada una de estas etiquetas pueden estar precedidas por ‘TIMEOUT_’ cuando un *timeout* ocurra mientras se espera a las respuestas ICP.

10. *Content-type*: Es el campo *content-type* de la cabecera de la respuesta HTTP.

2.1.2. PARTICIONADO Y CLASIFICACIÓN

Estas son las dos tareas que realiza la aplicación, es decir, trocea todas y cada una de las trazas y clasifica los campos que lo requieran, esto es, forma grupos con los campos cualitativos. Con más detalle este es el tratamiento que se da a los campos de cada una de las trazas:

- El *Tiemstamp*. No requiere de ningún tratamiento. Es escrito tal cual en el fichero ‘out0’
- *Elapsed time*. No requiere de ningún tratamiento. Es escrito tal cual en el fichero ‘out1’
- Dirección IP del cliente. Se quitan los puntos propios del formato IP y se escribe el resultado al fichero ‘out2’.
- Etiqueta de tratamiento y código HTTP. Este campo es a su vez dividido en dos. Por un lado la etiqueta de tratamiento se clasificara según la tabla 2.2.1 y escrito en el fichero ‘out3’, y por otro lado el código HTTP que será escrito en el fichero ‘out4’.

ETIQUETA	VALOR
TCP_HIT	1
TCP_MEM_HIT	2
TCP_MISS	3
TCP_REFRESH_HIT	4
TCP_REF_FAIL_HIT	5
TCP_REFRESH_MISS	6
TCP_CLIENT_REFRESH_MISS	7
TCP_IMS_HIT	8
TCP_IMS_MISS	9
TCP_SWAPFAIL_MISS	10
TCP_DENIED	11

Tabla 2.1. Asociación de valores a las etiquetas de tratamiento del dato.

- Tamaño. Se escribe al fichero ‘out5’ sin recibir ninguna modificación.
- Método de la petición. Es clasificado según la tabla 2.2.2 y escrita al fichero ‘out6’.

ETIQUETA	VALOR
GET	1
POST	2
HEAD	3

Tabla 2.2. Asociación de valores a las etiquetas de método de la petición.

- URL. No recibe ninguna modificación. Es escrito al fichero ‘out7’.
- Identidad del usuario. No recibe ninguna modificación. Es escrito al fichero ‘out8’.
- jerarquía de datos y nombre del host. Este campo es a su vez troceado en dos. Por un lado la jerarquía del dato que se clasifican según la tabla 2.2.3 y escrita al fichero ‘out9’, y por otro la dirección IP del host recibe el mismo trato que la del cliente y es escrita al fichero ‘out10’.

ETIQUETA	VALOR
NONE	1
DIRECT	2
FIRST_PARENT_MISS	3
SIBLING_HIT	4
PARENT_HIT	5
SINGLE_PARENT	6
CLOSEST_PARENT_MISS	7
CLOSEST_DIRECT	8
CACHE_DIGEST_HIT	9
CLOSEST_PARENT	10

Tabla 2.3. Asociación de valores para las distintas etiquetas de jerarquías de datos.

- *Content-type*. Es escrito tal cual al fichero ‘out11’.
- En el fichero ‘out12’ se escribe la extensión del objeto solicitado. Esto se realiza de la siguiente manera: primero intenta conseguirlo a partir de la URL, si aquí no estuviera presente, lo intentará conseguir a partir del content type. Por último si aquí no se encontrase, la extensión del fichero se considerará como desconocida. En la tabla 2.2.4 se establecen los rangos de valores para las distintas clases de objeto (ver [Apéndice A](#) para más de detalles).

CLASE	EJEMPLOS	VALORES
Imágenes	Jpg, gif, ...	[1..19]
Comprimidos	Zip, rar, ...	[37..46]
Aplicaciones	Exe, class, ...	[73..83]
Videos	Mpg, avi, ...	[109..122]

Sonidos	Wav, mp3, ...	[145..166]
Pag. estaticas	Html, htm, ...	[181..190]
Pag. dinamicas	Php, asp, ...	[217..229]
Textos	Txt, cfg, ...	[253..255]
HTTP	http	289
Otros	Dat, iso, ...	[325..342]
Scripts	Js, vbs, ...	[361..363]
Documentos	Xls, doc, ...	[397..420]
IR Cache	Irc	433
Presentaciones	Xsl, css	[469..470]
Desconocidos/Sin clasificar		[-2, -1]

Tabla 2.4. Clasificación de extensiones relaizada por la aplicación de pre-procesamiento.

2.1.3. EJEMPLO BÁSICO

Por ultimo, para un mejor entendimiento de la funcionalidad de esta aplicación, expondremos un ejemplo:

Dada la siguiente traza:

```
1080345620.363 23071 81.141.249.145 TCP_REFRESH_MISS/504 1649 GET http://quickweightloss.ws/tabel_back.gif - NONE/- text/html
```

que supondremos se encuentra en un fichero llamado pa.20040328 y situada en la ruta 'c:\proyecto\work\'.

Estos serian los ficheros que podrían ser generados:

- Out0: conteniendo el timestamp 1080345620.363
- Out1: conteniendo el lapso de tiempo 23071
- Out2: conteniendo el siguiente valor que representará a una IP: 81141249145
- Out3: conteniendo el valor 6
- Out4: conteniendo el código HTTP 504
- Out5: conteniendo el tamaño 1649
- Out6: conteniendo el valor 1
- Out7: conteniendo la URL http://quickweightloss.ws/tabel_back.gif
- Out8: conteniendo '-'
- Out9: conteniendo el valor 1

- Out10: conteniendo el valor ‘-‘
- Out11: conteniendo la cadena text/html
- Out12: conteniendo el valor 4 (valor asociado a las extensiones gif).

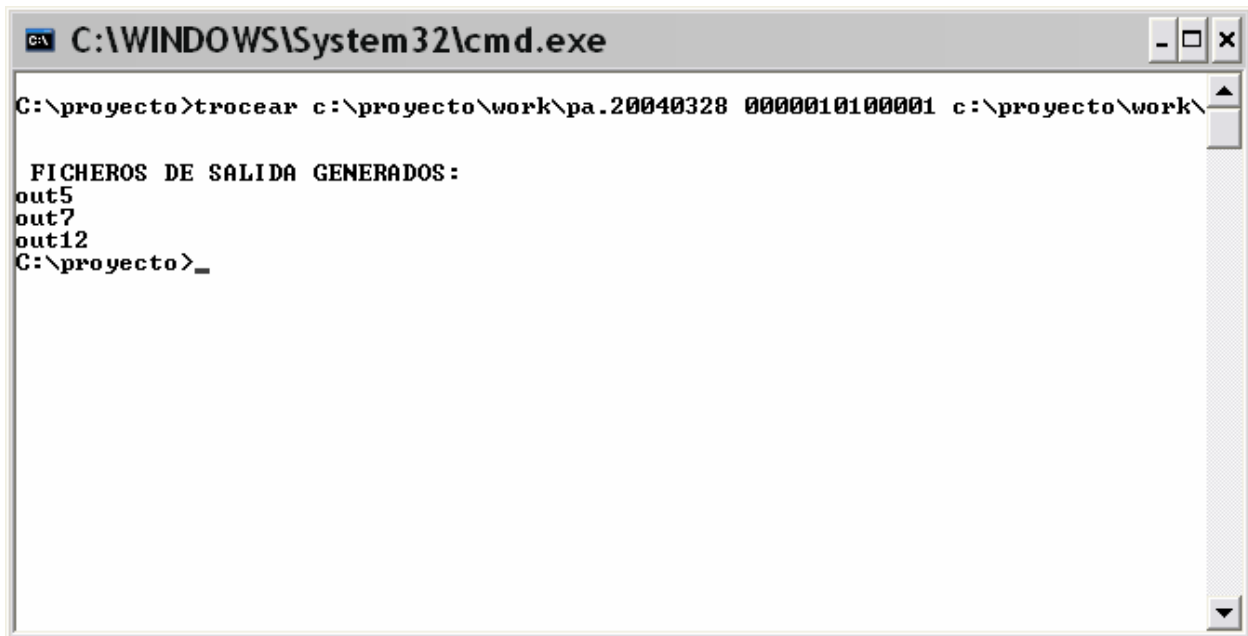


Figura 2.1. Ejemplo de uso de la aplicación de troceado.

En la figura 2.1 se observa la sintaxis desde la línea de comandos necesaria para la ejecución del ejemplo arriba expuesto. Los ficheros de salida que en este caso se han querido generar son únicamente out5, out7, y out12, los cuales serán ubicados en la misma ruta que el fichero de entrada, es decir: 'c:\proyecto\work\'.

2.1.4. FORMAS DE EJECUCIÓN

Existen dos formas de ejecutar esta aplicación:

1. Desde la línea de comandos, con la siguiente sintaxis:

trocear <fichero_trazas> <cadena_de_seleccion> <dir_trabajo>

Donde:

- <fichero_trazas> es la ruta y nombre del fichero de trazas
- <cadena_de_seleccion> es una cadena de 13 bits que indica que campo se quieren seleccionar

- <dir_trabajo> es el directorio donde se generaran los ficheros de salida.

Así por ejemplo

trocear c:\trazas\pa.sanitized-access.20040324 0000010100000 c:\proyecto\work generaría el fichero ‘out5’ con los tamaños y el fichero ‘out7’ conteniendo todas las URL, ambos ubicados en ‘c:\proyecto\work\’.

2. La otra forma es desde la aplicación principal que será vista con más detalle en el [apartado 2.5.2.11](#) junto con todas sus opciones del menú principal.

2.2. FRECUENCIAS URL

Su función es contar el número de peticiones a los distintos objetos, es decir, realiza el cálculo de las frecuencias con las que se solicitan todas y cada unas de las distintas URL presentes en un archivo de trazas. La entrada a esta aplicación es el archivo ‘out7’ generado por la aplicación de pre-procesamiento, es decir, las URL depuradas. Y su salida es la generación de dos archivos. Por un lado, un archivo conteniendo el listado de número de peticiones que ha recibido cada URL (en orden ascendente), por otro lado, el segundo archivo contendrá el listado de dichas URL.

2.2.1. EJEMPLO BÁSICO

Para una mejor comprensión de su funcionamiento, veamos un ejemplo:

Dado el siguiente fichero de entrada ‘out7’, generado por la aplicación de pre-procesamiento:

*http://quickweightloss.ws/tabel_back.gif
http://img.cdw.com/cdwg/homepage/thumbs/031504_spot3.gif
http://img.cdw.com/cdwg/homepage/thumbs/031504_spot4.gif
http://quickweightloss.ws/top.gif
http://quickweightloss.ws/left.gif
http://di.dell.com/images/global/general/borderedlayout/blue_bar_r.gif
http://www.elated.com/cgi-bin/imagekits/browse_images.pl?[1f5B74a9tN:f22JJS68liK]
http://w.w.w.ftipky.cz/video/trpите_stresem.mpg
http://w.w.w.ftipky.cz/video/trpите_stresem.mpg
http://w.w.w.ftipky.cz/video/trpите_stresem.mpg
http://zone.msn.com/NR/rdonlyres/D355CEEB-4A8C-4ABB-AFDB-8B73FE590EAE/18671/image004.gif
http://img.4chan.org/h/thumb/1080263175575s.jpg
http://www.spewdeevids.com/gallery16/index_files/squirting2_19.gif
http://swa.swavacations.com/images/head_destinations1.gif
http://www.elated.com/imagekits/0/thumbs/47.gif
http://www.elated.com/imagekits/0/thumbs/90.gif
http://www.elated.com/imagekits/1/thumbs/100.gif
http://www.elated.com/imagekits/1/thumbs/104.gif
http://www.elated.com/imagekits/1/thumbs/106.gif
http://www.elated.com/imagekits/1/thumbs/107.gif
http://charterone.com/images/nav_lock.gif
http://charterone.com/images/ehl.gif
http://zone.msn.com/NR/rdonlyres/D355CEEB-4A8C-4ABB-AFDB-8B73FE590EAE/18671/image004.gif
http://img.4chan.org/h/thumb/1080263175575s.jpg
http://www.elated.com/imagekits/0/thumbs/47.gif
http://www.elated.com/imagekits/0/thumbs/90.gif
http://www.elated.com/imagekits/1/thumbs/100.gif
http://www.elated.com/imagekits/1/thumbs/104.gif*

<http://www.elated.com/imagekits/1/thumbs/106.gif>
<http://www.elated.com/imagekits/1/thumbs/107.gif>

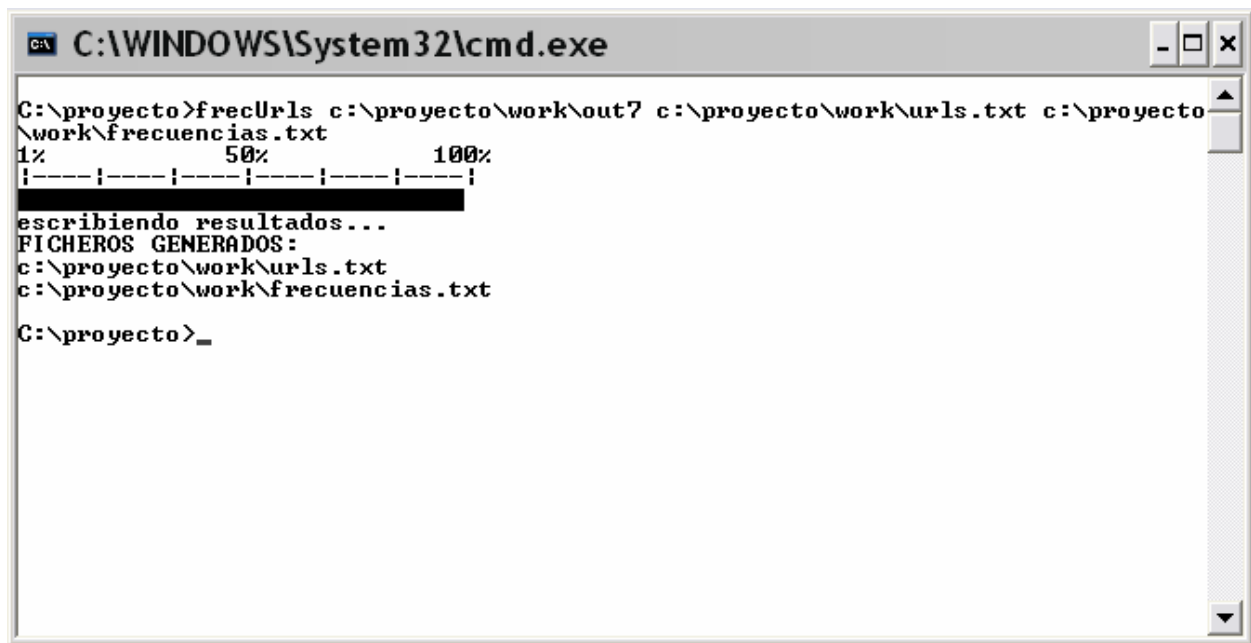
Tras la ejecución de esta aplicación obtendremos los siguientes resultados:

- Un fichero conteniendo la frecuencia de peticiones para cada una de las URL:

3
2
2
2
2
2
2
2
2
2
1
1
1
1
1
1
1
1
1
1
1

- Un fichero conteniendo la lista de estas URL.

w.w.ftipky.cz/video/trpite_stresem.mpg
www.elated.com/imagekits/1/thumbs/107.gif
img.4chan.org/h/thumb/1080263175575s.jpg
www.elated.com/imagekits/0/thumbs/47.gif
www.elated.com/imagekits/0/thumbs/90.gif
www.elated.com/imagekits/1/thumbs/100.gif
www.elated.com/imagekits/1/thumbs/104.gif
www.elated.com/imagekits/1/thumbs/106.gif
zone.msn.com/NR/rdonlyres/D355CEEB-4A8C-4ABB-AFDB-8B73FE590EAE/18671/image004.gif
img.cdw.com/cdwg/homepage/thumbs/031504_spot4.gif
www.spewdeevids.com/gallery16/index_files/squirting2_19.gif
swa.swavacations.com/images/head_destinations1.gif
quickweightloss.ws/top.gif
quickweightloss.ws/left.gif
di.dell.com/images/global/general/borderedlayout/blue_bar_r.gif
www.elated.com/cgi-bin/imagekits/browse_images.pl
quickweightloss.ws/tabel_back.gif
img.cdw.com/cdwg/homepage/thumbs/031504_spot3.gif
charterone.com/images/nav_lock.gif
charterone.com/images/ehl.gif



```
C:\WINDOWS\System32\cmd.exe
C:\proyecto>frecURLs c:\proyecto\work\out7 c:\proyecto\work\urls.txt c:\proyecto\work\frecuencias.txt
1%          50%          100%
|-----|-----|-----|-----|-----|
|-----|-----|-----|-----|-----|
escribiendo resultados...
FICHEROS GENERADOS:
c:\proyecto\work\urls.txt
c:\proyecto\work\frecuencias.txt
C:\proyecto>_
```

Figura 2.2. Ejemplo de ejecución de la aplicación de frecURL.exe.

En la figura 2.2 se observa la orden desde la línea de comandos necesaria para la ejecución del ejemplo anterior. Aquí los ficheros resultantes serán 'urls.txt' y 'frecuencias.txt' que serán ubicados en la ruta: 'c:\proyecto\work\'.

2.2.2. FORMAS DE EJECUCIÓN

Existen dos formas para ejecutar dicha aplicación:

1. Como cualquier aplicación desarrollada en C puede ejecutarse desde la línea de comandos con la siguiente sintaxis:

frecURL <fichero_entrada> <fichero_salida_URL> <fichero_salida_frec> [separador]

donde,

- <fichero_entrada> ruta y nombre del fichero 'out7' generado por la aplicación de pre-procesamiento, tal y como anteriormente se mencionaba
- <fichero_salida_URL> ruta y nombre del fichero donde se generará el listado de URL (ordenadas en orden ascendente de uso)
- <fichero_salida_frec> ruta y nombre del fichero donde se generara el listado de frecuencias de petición para cada una de las URL del fichero anterior.

El último argumento es opcional y será el separador que la URL usa para sus argumentos. Como norma general y por defecto '?'.
2. Otra forma de ejecutar esta aplicación es a través de la aplicación principal, mediante la opción '[Gráfica de las URL mas visitadas](#)' y eligiendo la opción 'Si' en la ventana modal mostrada a continuación.

2.3. ASIGNADOR DE CODIGOS HASH A URL

Esta aplicación nace con la idea de facilitar el procesamiento de las URL por parte de MatLab. Su función es la de asignar un código hash distinto a cada URL distinta. La entrada a esta aplicación será el archivo ‘out7’ generado por la aplicación de pre-procesamiento, y la salida serán dos archivos de texto. El primero de ellos contendrá, por cada línea, el par código hash – URL separados por un espacio en blanco, y el segundo archivo será igual al archivo de entrada (‘out7’) pero usando los ya asignados códigos hash.

Esta aplicación será más tarde utilizada por la herramienta principal para el cálculo de las distancias entre peticiones que se verá a lo largo de los apartados 2.5.2.8, 2.5.2.9, y 2.5.2.10.

2.3.1. EJEMPLO BÁSICO

Para comprender mejor la funcionalidad de esta aplicación veremos a continuación como se procesaría un archivo de entrada muy básico:

```
http://quickweightloss.ws/tabel_back.gif
http://img.cdw.com/cdwg/homepage/thumbs/031504_spot3.gif
http://img.cdw.com/cdwg/homepage/thumbs/031504_spot4.gif
http://quickweightloss.ws/top.gif
http://quickweightloss.ws/left.gif
http://di.dell.com/images/global/general/borderedlayout/blue_bar_r.gif
http://www.elated.com/cgi-bin/imagekits/browse_images.pl?[1f5B74a9tN:f22JJS68liK]
http://w.w.w.ftipky.cz/video/trpite_stresem.mpg
http://w.w.w.ftipky.cz/video/trpite_stresem.mpg
http://w.w.w.ftipky.cz/video/trpite_stresem.mpg
http://zone.msn.com/NR/rdonlyres/D355CEEB-4A8C-4ABB-AFDB-8B73FE590EAE/18671/image004.gif
http://img.4chan.org/h/thumb/1080263175575s.jpg
http://www.spewdeevids.com/gallery16/index_files/squirting2_19.gif
http://swa.swavacations.com/images/head_destinations1.gif
http://www.elated.com/imagekits/0/thumbs/47.gif
http://www.elated.com/imagekits/0/thumbs/90.gif
http://www.elated.com/imagekits/1/thumbs/100.gif
http://www.elated.com/imagekits/1/thumbs/104.gif
http://www.elated.com/imagekits/1/thumbs/106.gif
http://www.elated.com/imagekits/1/thumbs/107.gif
http://charterone.com/images/nav_lock.gif
http://charterone.com/images/ehl.gif
http://zone.msn.com/NR/rdonlyres/D355CEEB-4A8C-4ABB-AFDB-8B73FE590EAE/18671/image004.gif
http://img.4chan.org/h/thumb/1080263175575s.jpg
http://www.elated.com/imagekits/0/thumbs/47.gif
http://www.elated.com/imagekits/0/thumbs/90.gif
http://www.elated.com/imagekits/1/thumbs/100.gif
http://www.elated.com/imagekits/1/thumbs/104.gif
```

<http://www.elated.com/imagekits/1/thumbs/106.gif>
<http://www.elated.com/imagekits/1/thumbs/107.gif>

Como resultado a esta entrada, la aplicación generará dos archivos:

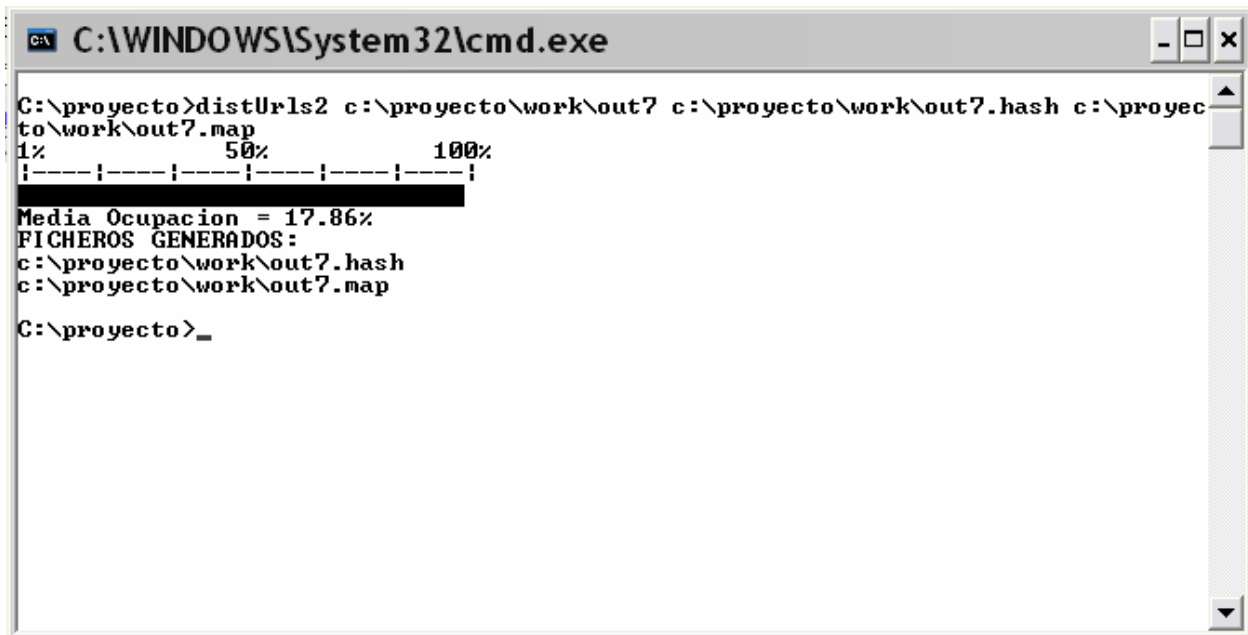
- Por un lado el par código hash – URL, que informa al usuario de la correspondencia establecida por la aplicación:

40001 img.cdw.com/cdwg/homepage/thumbs/031504_spot3.gif
40002 img.cdw.com/cdwg/homepage/thumbs/031504_spot4.gif
80001 di.dell.com/images/global/general/borderedlayout/blue_bar_r.gif
80002 www.elated.com/cgi-bin/imagekits/browse_images.pl
80003 www.elated.com/imagekits/0/thumbs/47.gif
80004 www.elated.com/imagekits/0/thumbs/90.gif
80005 www.elated.com/imagekits/1/thumbs/100.gif
80006 www.elated.com/imagekits/1/thumbs/104.gif
80007 www.elated.com/imagekits/1/thumbs/106.gif
80008 www.elated.com/imagekits/1/thumbs/107.gif
200001 quickweightloss.ws/tabel_back.gif
200002 quickweightloss.ws/top.gif
200003 quickweightloss.ws/left.gif
360001 www.spewdeevs.com/gallery16/index_files/squirting2_19.gif
360002 swa.swavacations.com/images/head_destinations1.gif
380001 charterone.com/images/nav_lock.gif
380002 charterone.com/images/ehl.gif
440001 w.w.w.fiipky.cz/video/trpite_stresem.mpg
520001 zone.msn.com/NR/rdonlyres/D355CEEB-4A8C-4ABB-AFDB-8B73FE590EAE/18671/image004.gif
520002 img.4chan.org/h/thumb/1080263175575s.jpg

- Por otro lado el mismo archivo de entrada pero sustituyendo cada una de las URL por su código hash:

200001
40001
40002
200002
200003
80001
80002
440001
440001
440001
520001
520002
360001
360002
80003
80004
80005
80006
80007
80008

380001
380002
520001
520002
80003
80004
80005
80006
80007
80008



```
C:\WINDOWS\System32\cmd.exe
C:\proyecto>distURLs2 c:\proyecto\work\out7 c:\proyecto\work\out7.hash c:\proyecto\work\out7.map
1%
Media Ocupacion = 17.86%
FICHEROS GENERADOS:
c:\proyecto\work\out7.hash
c:\proyecto\work\out7.map
C:\proyecto>_
```

Figura 2.3. Ejemplo de ejecución de la aplicación distURL2.exe.

En la figura 2.3 se puede observar la sintaxis desde la línea de comandos necesaria para la ejecución del ejemplo arriba expuesto. En este los ficheros generados son llamados ‘out7.hash’ y ‘out7.map’, ambos en ‘c:\proyecto\work\’.

2.3.2. FORMAS DE EJECUCIÓN

Como ya viene siendo habitual en todas las aplicaciones del proyecto, esta también tiene dos formas de ser invocada:

1. Desde la línea de comandos con la siguiente sintaxis:

distURL2 <fichero_URL> <fichero_salida_hash> <fichero_salida_map>

donde:

- <fichero_URL> es la ruta y nombre del fichero ‘out7’ generado por la aplicación de pre-procesamiento

- <fichero_salida_hash> es la ruta y nombre del fichero que queremos usar para escribir la traducción de <fichero_URL> a códigos hash
 - <fichero_salida_map> es la ruta y nombre del fichero donde se establece la correspondencia utilizada para dicha traducción.
2. La siguiente forma es mediante la aplicación principal, con cualquiera de las opciones siguientes: [Distancias entre peticiones de un objeto](#), [Tamaño entre peticiones](#), o [Tiempo entre peticiones](#). En los tres casos esta aplicación servirá como pre-procesamiento a otros cálculos. Estas opciones serán comentadas con mas detalle en el [apartado 2.5](#).

2.4. SIMULADOR DE PROXY-CACHE

Se trata de una aplicación desarrollada en java que simula la cache de un proxy, proporcionando tras la simulación la tasa de hits y hit bytes resultantes.

A lo largo de este capítulo veremos las distintas formas de simulación, el ajuste de los parámetros de entrada, y la forma de ejecución.

2.4.1. POLÍTICAS DE REEMPLAZO

Es el criterio aplicado cuando es necesario almacenar un nuevo objeto solicitado y el tamaño de este es mayor que el espacio libre en cache. Se elige entonces uno o varios objetos de cache para liberar espacio. Las políticas de reemplazo intentan entonces causar el mínimo impacto posible en el rendimiento de la cache, desalojando aquellos objetos que podrían ser los menos usados en futuro.

Esta aplicación implementa las siguientes políticas de reemplazo:

- LRU (Least Recently Used): consiste en reemplazar aquel objeto que llego en el menor instante de tiempo.
- LFU (Least Frequently Used): consiste en reemplazar el objeto menos usado. En este caso por tanto cada objeto llevara asociado un contador con el número de veces que se ha solicitado. Es posible que existiera más de un objeto candidatos para el reemplazo. En este caso se aplicará una política LRU.
- Existe otra posibilidad que es la de mezclar ambas políticas. Para este caso, una parte de la cache implementaría una de las dos políticas de reemplazo arriba expuestas y almacenaría un determinado tipo de objeto. Mientras, la otra parte almacenaría otro tipo de objetos y aplicaría otra política de reemplazo.

Por ejemplo: imaginemos una cache de tamaño T , en la que una parte de la misma de tamaño t_1 con $0 < t_1 < T$ almacenaría solamente imágenes y seguiría una política de reemplazo LRU. El resto $(T - t_1)$ almacenaría solamente objetos de sonido, siguiendo un reemplazo LFU.

2.4.2. SIMULACIÓN SIMPLE Y POR LOTES

Existe dos clases con métodos main y que por tanto son puntos de entrada a la aplicación: Simulación y SimulacionBatch.

La primera solo ejecuta una simulación dada, es decir, un tamaño, un tipo de objeto, y con un tipo de reemplazo. Por otro lado SimulacionBatch ofrece los resultados de caches LRU y LFU, con distintas capacidades, y para varias clases de objeto. De esta manera por ejemplo, podemos observar las distintas curvas representando el rendimiento de las caches dependiendo de su tamaño y clases de objetos. Es por tanto el equivalente a lanzar varias veces la clase Simulación variando tamaño de cache, tipo de objeto a almacenar, y tipo de cache. Ambas clases toman sus parámetros de entrada desde un fichero de propiedades.

2.4.3. ARCHIVO DE PROPIEDADES

Es el fichero ‘Proyecto.properties’ situado en la carpeta ComponentsProperties. En este fichero ajustamos los distintos parámetros de entrada a la aplicación. Estas propiedades nos permiten configurar:

- la ubicación de los distintos ficheros (Ej.: fichero de timestamp, de URL,...). Tanto de entrada como de salida.
- tamaño de la cache a simular
- tipo de objetos a tratar
- política de reemplazo (LRU/LFU).

A continuación, en la Tabla 5.3.1 se muestran las propiedades que necesitan ser configuradas, bien para ejecutar una sola simulación o bien en el caso de ejecutar varias simulaciones por lotes:

Nombre	Múltiples valores	Ajustar para	Descripción
PROJECT_SIMULATION_SIZECACHE	No	Simulación	Tamaño de la cache a simular
PROJECT_SIMULACIONBATCH_SIZECACHE	Si	SimulacionBatch	Tamaño máximo de la cache
PROJECT_SIMULACIONBATCH_MAXPERCENT	No	SimulacionBatch	Porcentaje del tamaño máximo con el que se inicia la

PROJECT_SIMULACIONBATCH_MINPERCENT	No	SimulacionBatch	simulación Porcentaje del tamaño mínimo con el que se finaliza la simulación
PROJECT_SIMULACIONBATCH_DECPCENT	No	SimulacionBatch	Decremento del tamaño de la cache entre sucesivas simulaciones
PROJECT_SIMULATION_TYPECACHE	No	Simulación	Tipo de cache a simular (LRU/LFU)
PROJECT_HASHURL_INPUTFILELOCATION_URL	No	HashURL	Ruta completa del fichero de URL
PROJECT_HASHURL_OUTPUTLOCATION	No	HashURL	Ruta completa del fichero de salida
PROJECT_SIMULATION_PATH_INPUTFILELOCATION_IDENTIFIERS	No	Simulación / SimulacionBatch	Ruta completa del fichero hash de URL
PROJECT_SIMULATION_PATH_INPUTFILELOCATION_SIZES	No	Simulación / SimulacionBatch	Ruta completa del fichero de tamaños
PROJECT_SIMULATION_PATH_INPUTFILELOCATION_TIMESTAMPS	No	Simulación / SimulacionBatch	Ruta completa del fichero de timestamps
PROJECT_SIMULATION_PATH_INPUTFILELOCATION_TYPES	No	Simulacion / SimulacionBatch	Ruta completa del fichero de tipos de objeto
PROJECT_SIMULACIONBATCH_OUTPUTFILE	No	SimulacionBatch	Ruta completa del fichero CSV de salida
PROJECT_SIMULATION_EXTENSIONS_DINPAGE_LOW	No	Simulacion / SimulacionBatch	Mínimo valor para los tipo de paginas dinámicas
PROJECT_SIMULATION_EXTENSIONS_DINPAGE_HIGH	No	Simulacion / SimulacionBatch	Máximo valor para el tipo de paginas dinámicas
PROJECT_SIMULATION_EXTENSIONS_TOSAVE_LOW	No	Simulacion	Mínimo valor para los tipo a utilizar en la simulación
PROJECT_SIMULATION_EXTENSIONS_TOSAVE_HIGH	No	Simulacion	Máximo valor para los tipo a utilizar en la simulación
PROJECT_SIMULACIONBATCH_EXTENSIONS_TOSAVE_LOW	Si	SimulacionBatch	Mínimo valor para los tipo a utilizar en la simulación
PROJECT_SIMULACIONBATCH_EXTENSIONS_TOSAVE_HIGH	Si	SimulacionBatch	Máximo valor para los tipo a utilizar en la simulación

Tabla 2.5. Listado de propiedades que deben ser configuradas.

2.4.3.1. Ajuste del tipo de cache

Para el caso de ejecutar una sola simulación, es decir invocar el método main de la clase Simulación, es posible ajustar el tipo de cache a simular. Esta propiedad es PROJECT_SIMULATION_TYPECACHE, y sus posibles valores son LRU o LFU. En el caso

de ejecutar varias simulaciones por lotes, es decir invocamos el método main de la clase SimulacionBatch, el tipo de cache no necesita ser especificado sino que para cada tamaño y para cada tipo de objeto a estudiar, la clase SimulacionBatch simulará alternativamente las políticas LRU y LFU.

2.4.3.2. Ajuste del tamaño de cache

Para el caso de querer realizar una sola simulación, será necesario ajustar la propiedad PROJECT_SIMULATION_SIZECACHE para especificar la capacidad de la cache, donde las unidades se expresaran en bytes.

En el caso de querer realizar simulaciones por lotes, ajustaremos la propiedad PROJECT_SIMULACIONBATCH_SIZECACHE. Además, las propiedades PROJECT_SIMULACIONBATCH_MAXPERCENT, PROJECT_SIMULACIONBATCH_MINPERCENT, y PROJECT_SIMULACIONBATCH_DECPERCENT permiten ir decrementando el tamaño inicial para poder estudiar la evolución del rendimiento con respecto al tamaño. En pseudo código quedaría así:

```
PARA P= PROJECT_SIMULACIONBATCH_MAXPERCENT HASTA  
PROJECT_SIMULACIONBATCH_MINPERCENT PASO -  
PROJECT_SIMULACIONBATCH_DECPERCENT HACER  
    Simular Cache con tamaño PROJECT_SIMULACIONBATCH_SIZECACHE*(P/100)  
FINPARA
```

2.4.3.3. Ajuste de la ubicación de los ficheros

Los ficheros necesarios para la ejecución de una simulación son:

1. Fichero hash de URL, bien obtenido a través de la aplicación explicada en el [apartado 2.3](#), o bien mediante el método main de la clase HashURL y ajustando las propiedades necesarias para esta clase tal y como muestra la tabla 2.5
2. Ficheros de timestamp, de tamaños, y de tipos, es decir, el fichero 'out0', 'out5', y 'out12', generados por la aplicación de troceado expuesto en el [apartado 2.1](#).

La ruta completa y nombre de estos ficheros deberá ser ajustada en el fichero de propiedades tal y como se expone en la tabla 2.5.

2.4.3.4. Ajuste de los tipos de objetos a almacenar

Con las propiedades `PROJECT_SIMULATION_EXTENSIONS_TOSAVE_LOW` y `PROJECT_SIMULATION_EXTENSIONS_TOSAVE_HIGH` es posible determinar qué tipo de objetos se almacenaran en cache. En el apéndice de clasificación de extensiones se exponen los valores máximos y mínimos para cada tipo de objeto. De la misma manera, las propiedades `PROJECT_SIMULACIONBATCH_EXTENSIONS_TOSAVE_LOW` y `PROJECT_SIMULACIONBATCH_EXTENSIONS_TOSAVE_HIGH` son necesarias para el caso de varias simulaciones por lotes (SimulacionBatch). En este caso es posible ajustar varios tipos para sucesivas simulaciones, separándolos por el carácter '&'.

2.4.4. EJEMPLO BÁSICO

Supongamos por ejemplo que quisiéramos simular y comparar el rendimiento de un proxy-cache cuando ésta almacena solo objetos de tipo imagen y cuando almacena solo objetos de tipo video. Además deseamos calcular el rendimiento de ambas políticas cuando vamos decrementando el tamaño de la cache en un 10%.

Veamos entonces, a partir de este ejemplo, la configuración del fichero de propiedades. Para empezar, y puesto que se desean realizar varias simulaciones, solo será necesario ajustar aquellas propiedades que tengan el prefijo '`PROJECT_SIMULACIONBATCH_`', que son las únicas propiedades que afectan a la parte de la aplicación que vamos a lanzar.

Ajustamos en primer lugar el tamaño máximo inicial que vamos a dar a la cache, por ejemplo 99.999 bytes:

```
PROJECT_SIMULACIONBATCH_SIZECACHE=99999
```

Para que la cache solo almacene objetos de tipo imagen realizamos el siguiente ajuste:

```
PROJECT_SIMULACIONBATCH_EXTENSIONS_TOSAVE_LOW=1
```

```
PROJECT_SIMULACIONBATCH_EXTENSIONS_TOSAVE_HIGH=17
```

puesto que el rango de valores para los objetos de tipo imagen es del 1 al 17 (ver Apéndice A).

Para realizar varias simulaciones donde el tamaño se vaya decrementado en un 10%, realizamos los siguientes tres ajustes:

PROJECT_SIMULACIONBATCH_MAXPERCENT=100

PROJECT_SIMULACIONBATCH_MINPERCENT=10

PROJECT_SIMULACIONBATCH_DECPERCENT=10

donde la primera de ellas especifica que partimos del 100% del tamaño antes ajustado, la segunda que la aplicación finalice cuando llegue al 10% de dicho tamaño, y la tercera que se decremente el tamaño en un 10% en cada nueva simulación.

Además de todo esto será imprescindible la ubicación de los ficheros ‘out0’, ‘out5’, ‘out7’, y ‘out12’, generados por la aplicación de pre-procesamiento (explicada en el [capítulo 1](#)).

Suponiendo que la ruta de estos sea ‘C:\Proyecto’, tendremos:

PROJECT_SIMULATION_PATH_INPUTFILELOCATION_SIZES=C:/Proyecto/out5

PROJECT_SIMULATION_PATH_INPUTFILELOCATION_TIMESTAMPS=C:/Proyecto/out0

PROJECT_SIMULATION_PATH_INPUTFILELOCATION_TYPES=C:/Proyecto/out12

PROJECT_HASHURL_INPUTFILELOCATION_URL=C:/Proyecto/out7

Por ultimo la propiedad que determina el nombre del fichero generado por la aplicación de HashURL y la propiedad del fichero de entrada de códigos hash de URL de la aplicación, deberían coincidir:

PROJECT_HASHURL_OUTPUTLOCATION=C:/Proyecto/ejemplo

PROJECT_HASHURL_OUTPUTFILEEXTENSION=.hash

PROJECT_SIMULATION_PATH_INPUTFILELOCATION_IDENTIFIERS=C:/Proyecto/ejemplo.hash

2.4.5. FORMAS DE EJECUCIÓN

La forma más sencilla y cómoda es utilizando los ficheros scripts (‘run_simulacion.bat’ para el caso de querer ejecutar una sola simulación, y ‘run_simulacionBatch.bat’ si deseamos ejecutar varias simulaciones sucesivas). Ambos necesitan ser editados para ajustar el correcto valor de las variables PROJECT_HOME y LOG4J_HOME que dependerán de la ubicación de la aplicación y del fichero log4j-1.2.8.jar respectivamente. Ambos scripts llaman previamente a la clase HashURL para la obtención del fichero de hash de URL, necesitado para la ejecución de las simulaciones, facilitando de esta manera la invocación. Por supuesto, y como paso previo a la utilización a estos scripts será necesario ajustar el fichero de propiedades de la aplicación.

Por último comentar la existencia del fichero log donde se escribe información durante la ejecución de la aplicación que resulta útil, bien para conocer el grado de avance de la simulación,

bien para conocer cualquier posible fallo en la misma. En el fichero de propiedades (log4j.properties) pueden configurarse algunas características de este fichero, tal como tamaño máximo, ubicación,...

2.4.6. JAVADOC DE LA APLICACIÓN

El Javadoc de esta aplicación se encuentra en:

[.\Simulacion\doc\index.html](#)

2.5. PROGRAMA PRINCIPAL

Desde esta aplicación, desarrollada en MatLab, es posible invocar cualquiera de las aplicaciones comentadas en capítulos anteriores, bien directamente a través del menú principal, bien indirectamente si alguna opción necesita de ella. A lo largo de este capítulo navegaremos por todas sus opciones explicando con detalle cada una de ellas.

2.5.1. INICIO

El primer paso será asegurarnos que todos los ejecutables y scripts de lanzamiento, externos a esta aplicación, se encuentran situados en el directorio 'c:\Proyecto', estos son:

- frecURL.exe
- distURL2.exe
- trocear.exe
- run_simulacion.bat
- run_simulacionBatch.bat

Además estos dos últimos deben haber sido configurados previamente, tal y como especifica el [apartado 5.5](#).

En segundo lugar es necesario que MatLab encuentre los ficheros .m necesarios para la ejecución de esta aplicación. Elegimos por tanto, el directorio C:\Proyecto\matlab en el navegador de carpetas de Matlab, tal y como muestra la figura 6.1.1

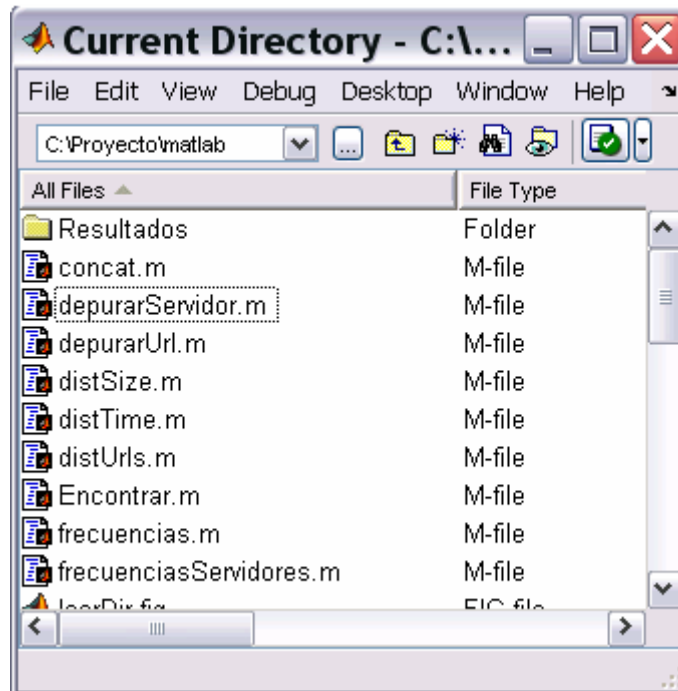


Figura 2.4. Elección del directorio de trabajo en MatLab.

Otra opción es incluir esta carpeta en el path de Matlab, mediante la opción File -> Set path... De esta forma no es necesario elegir esta carpeta cada vez que iniciemos Matlab.

Para iniciar la aplicación es necesario teclear 'proyecto' en la línea de comandos de Matlab, que hará visible el menú principal. En este menú podemos encontrar todas las opciones necesarias para poder generar cualquier resultado que se muestre en los apartados posteriores.

2.5.2. OPCIONES DEL MENU PRINCIPAL

Para cada una de las opciones, que componen el menú principal, expondremos una breve descripción de su cometido, las entradas que necesita (argumentos, ficheros,...) y las salidas que genera.

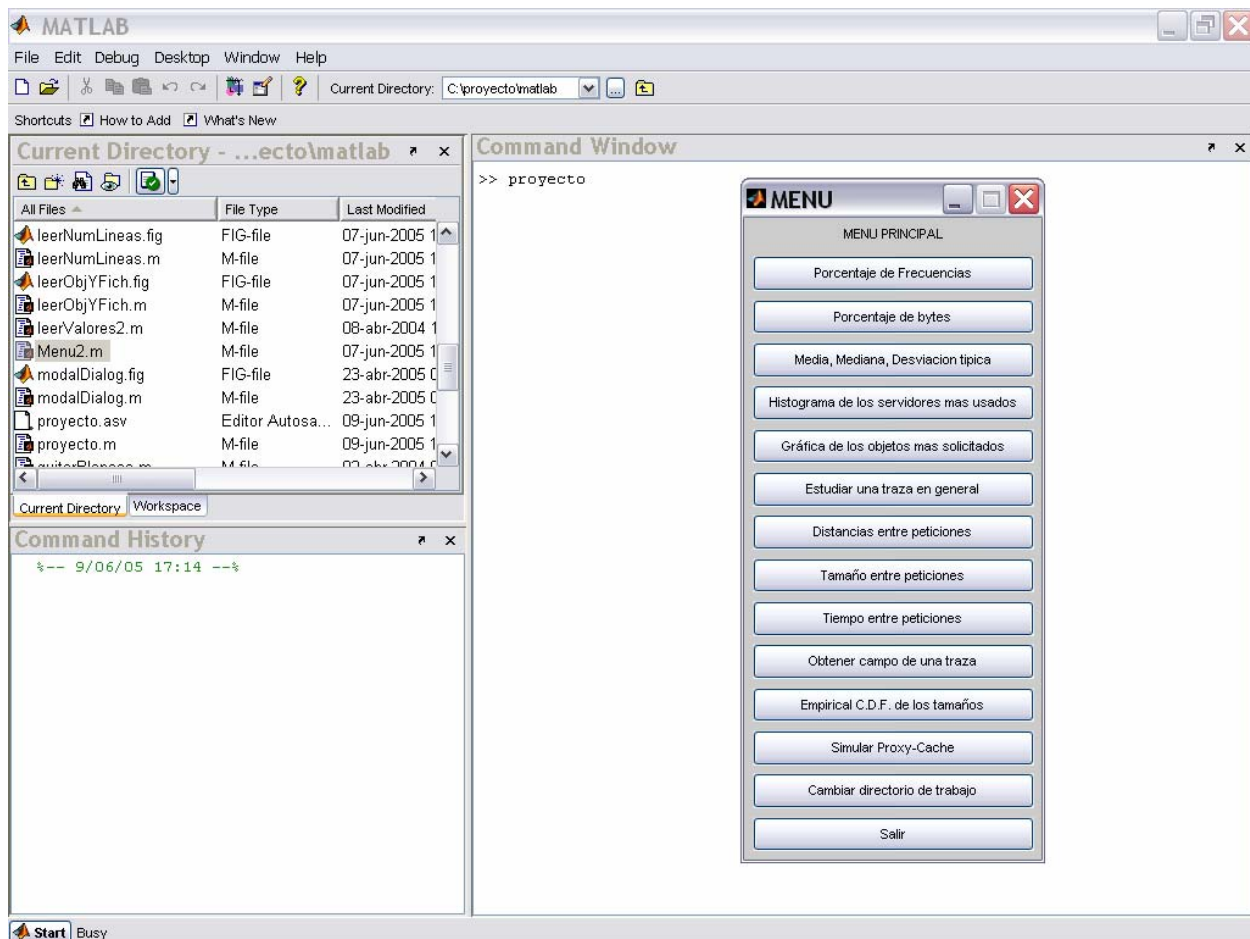


Figura 2.5. Menú principal de la aplicación.

2.5.2.1. Cambiar el directorio de trabajo

- Descripción: El directorio de trabajo es aquel al que apunta la aplicación. La ruta de los ficheros de entrada (como por ejemplo los ficheros de trazas), y la ubicación que tendrán los ficheros de salida vienen determinados por él. Una incorrecta elección en el mismo provocaría que los datos de entrada no se encuentren, dando lugar a un error parecido al siguiente:

Error using ==> load

Unable to read file c:\Proyecto\out12: No such file or directory.

- Entrada: ruta absoluta del fichero de trabajo.
- Salida: ninguna

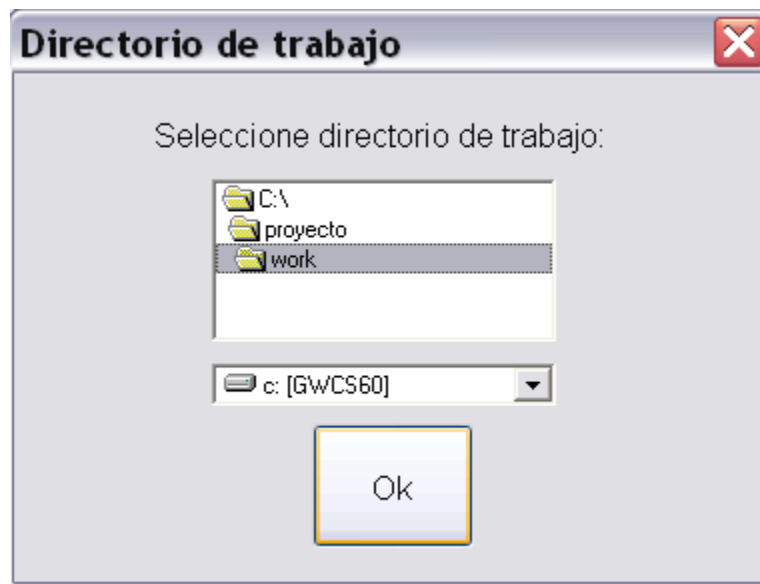


Figura 2.6. Elección del directorio de trabajo.

2.5.2.2. Porcentaje de frecuencias

- Descripción: calcula el porcentaje de veces que se han solicitado clases de objetos (imágenes, videos,...), o tipos en concreto (gif, jpeg, avi,...).
- Entrada: en la figura 6.2.3 se muestra la ventana que se abre tras elegir esta opción y en la que se introducirán el rango de tipos que se desea estudiar. En el apéndice A, la tabla de 'clasificación de extensiones' establece la correspondencia entre valores y tipos de objeto. Es necesario también, el fichero de tipos ('out12') generado por la aplicación de troceado que deberá estar situado en el directorio de trabajo.
- Salida: se genera por pantalla el porcentaje de veces que se ha solicitado la determinada clase o tipo de objeto con respecto al total de peticiones.

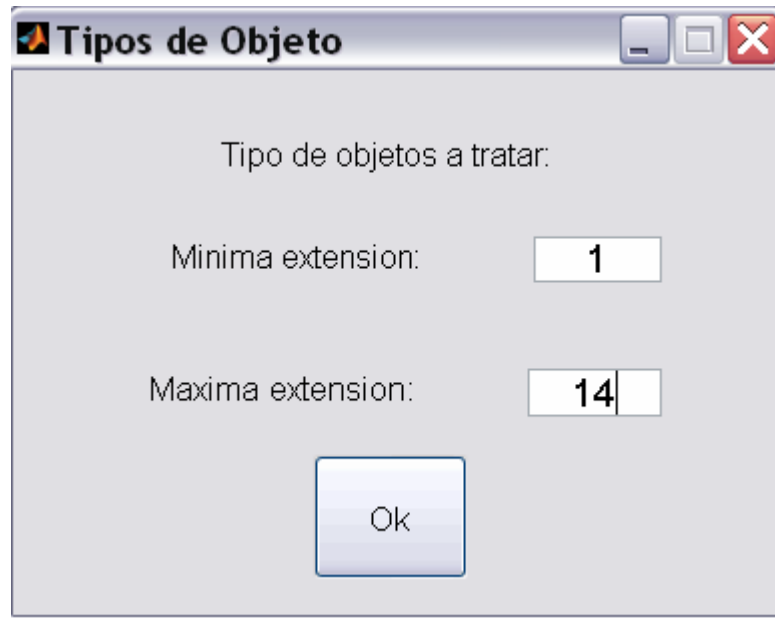


Figura 2.7. Ventana modal para la elección de los tipos de objetos a procesar.

2.5.2.3. Porcentaje de bytes

- Descripción: calcula el porcentaje de bytes que aportan al tráfico clases o tipos de objetos en particular.
- Entrada: rango de tipos al que se desea calcular el porcentaje de bytes, rango de códigos HTTP para los que solo se realizará el cálculo. Así mismo los ficheros de tipos, de tamaños, y de códigos HTTP (ficheros 'out12', 'out5', y 'out4' respectivamente, generados por la aplicación de troceado de trazas) deberán estar situados en el directorio de trabajo.
- Salida: muestra en la ventana de comandos de matlab el porcentaje de bytes con respecto al total de un determinado rango de objetos (y que obtuvieron en su solicitud un código de respuesta comprendido en el rango pasado como argumento).

2.5.2.4. Media, mediana, y desviación típica

- Descripción: calcula la media, mediana, y desviación típica para los tamaños de una determinada clase de objetos o tipos en particular.

- Entrada: al igual que en la opción del menú anterior es necesario introducir el rango de tipos para los que se desea realizar el cálculo y el rango de códigos HTTP. También será necesario que los ficheros 'out4', 'out5', 'out12' hayan sido generados previamente por la aplicación de pre-procesado y situados en el directorio de trabajo.
- Salida: muestra en la ventana de comandos de matlab la media, mediana, y desviación típica de un determinado rango de objetos (y que obtuvieron en su solicitud un código de respuesta comprendido en el rango pasado como argumento).



Figura 2.8. Ventana modal para la elección del rango de códigos HTTP.

2.5.2.5. Histograma de los servidores más usados

- Descripción: calcula la lista de los servidores más usados y los representa en un histograma. Puesto que el calculo es complejo, el usuario puede desear decrementar el numero de trazas que se desean tener en cuenta a la hora de elaborar dicha lista (tal y como muestra la figura 6.2.5). La ventana mostrará por defecto el número de trazas que posee el archivo.
- Entrada: fichero 'out7' que deberá estar situado en el directorio de trabajo. Además el fichero 'out1' será utilizado para hacer una estimación del número de trazas que posee el archivo.

- Salida:
 - lista con los servidores más usados ordenada ascendentemente.
 - histograma de la frecuencia de peticiones que recibió cada servidor.

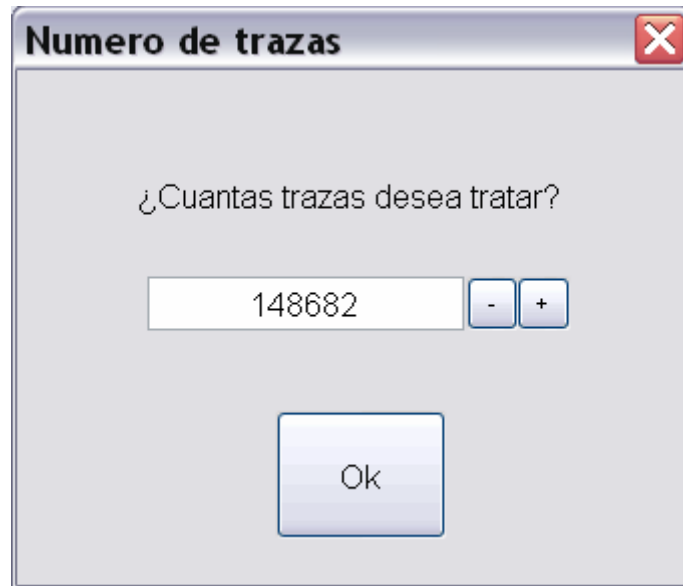


Figura 2.9. Selección del número de trazas a procesar.

2.5.2.6. Gráfica de las URL más visitadas

- Descripción: Invoca a la aplicación de frecuencias de URL expuesta en el [capítulo 3](#), es decir, calcula la frecuencia con la que se ha solicitado cada uno de los distintos objetos.
- Entrada: fichero 'out7' generado por la aplicación de pre-procesamiento situado en el directorio de trabajo.
- Salida:
 - fichero 'urls.txt' con la lista de URL ordenada según la frecuencia con la que se han solicitado
 - fichero 'frecs.txt' con el listado de frecuencias que ha recibido cada URL ordenado ascendentemente
 - gráfica de escala logarítmica, en ambos ejes, representando al fichero 'frecs.txt'

2.5.2.7. Estudiar una traza en general

- Descripción: realiza un estudio general de un fichero de trazas al completo.
- Entrada: ficheros 'out1', 'out4', 'out5', 'out6' generados por el programa de pre-procesamiento.
- Salida:
 - número de peticiones en total
 - porcentaje de uso para cada uno de los métodos (GET, POST, y HEAD).
 - Porcentaje de aparición de cada uno de los códigos HTTP (1xx, 2xx, 3xx, 4xx, y 5xx).
 - Media, mediana, y desviación típica de los tamaños y de los tiempos de respuesta.

2.5.2.8. Distancia entre peticiones a un objeto

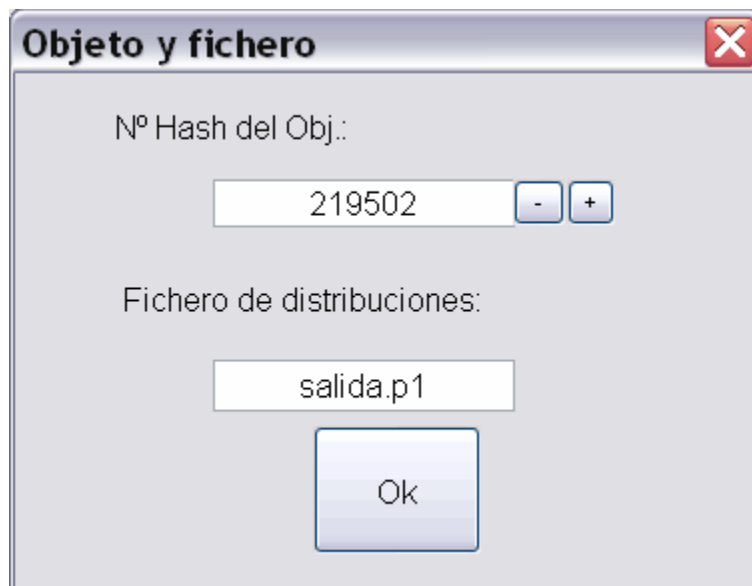
- Descripción: para peticiones sucesivas de un objeto dado, calcula el vector que contiene el espacio que existe entre ellas.
- Entrada: número hash de objeto a estudiar y fichero hash de URL en el directorio de trabajo.
- Salida:
 - vector con las distancias
 - gráfica del vector o de su acumulado

2.5.2.9. Tamaño entre peticiones a un objeto

- Descripción: para peticiones sucesivas de un objeto dado, calcula el vector que contiene la suma de los tamaños de los objetos que existen entre ellas.
- Entrada: número hash de objeto a estudiar, y ficheros hash de URL y de tamaños ('out5') en el directorio de trabajo.
- Salida:
 - vector con la suma de tamaños entre peticiones
 - gráfica del vector o de su acumulado

2.5.2.10. Tiempo entre peticiones a un objeto

- Descripción: para peticiones sucesivas de un objeto dado, calcula el vector que contiene la suma de los tiempos de los objetos que existen entre ellas.
- Entrada: número hash de objeto a estudiar, y ficheros hash de URL y de tamaños ('out0') en el directorio de trabajo.
- Salida:
 - vector con la suma de tiempos entre peticiones
 - gráfica del vector o de su acumulado



The image shows a Windows-style dialog box with the title "Objeto y fichero" and a red close button in the top right corner. Inside the dialog, there are two input fields. The first is labeled "Nº Hash del Obj.:" and contains the text "219502". To the right of this field are two small buttons, one with a minus sign "-" and one with a plus sign "+". The second input field is labeled "Fichero de distribuciones:" and contains the text "salida.p1". Below these fields is a single "Ok" button.

Figura 2.10. Elección el numero hash de objeto y fichero a estudiar.

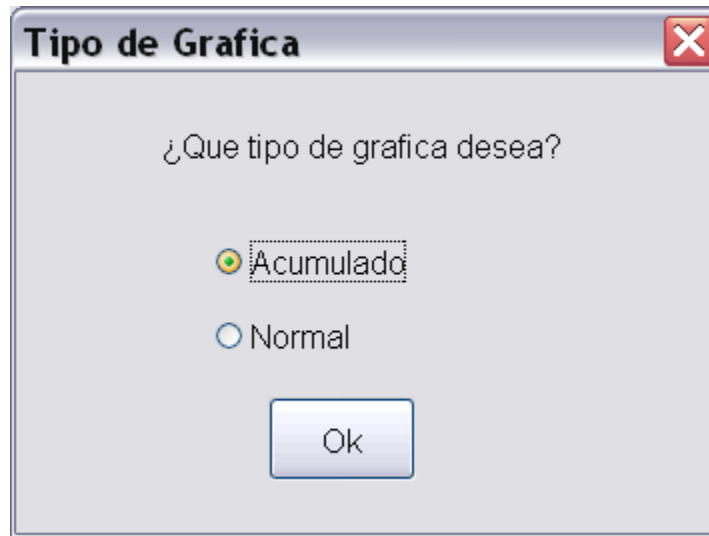


Figura 2.11. Elección del tipo de grafica a mostrar.

2.5.2.11. Obtención del campo de una traza

- Descripción: esta opción llama a la aplicación de pre-procesamiento, explicada en el [capítulo 2](#), para obtener el campo de una traza.
- Entrada:
 - nombre del fichero de trazas
 - campos a obtener
 - fichero de trazas y aplicación de pre-procesado ('trocear.exe') en el directorio de trabajo.
- Salida:
 - ficheros a los que se han copiado los campos de la traza seleccionados.

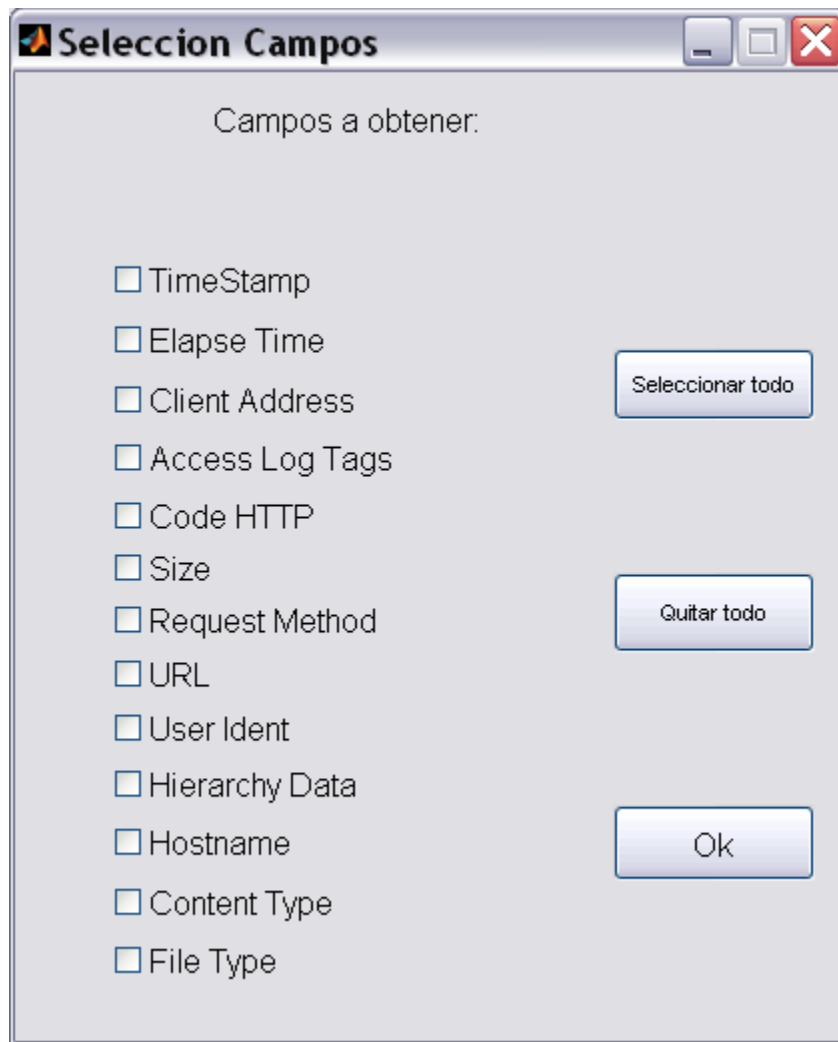


Figura 2.12. Elección de los campos de traza a obtener.

2.5.2.12. Función de distribución de los tamaños

- Descripción: representa gráficamente la función de distribución de los tamaños acumulados
- Entrada: fichero 'out5' generado por la aplicación de pre-procesamiento, y situado en el directorio de trabajo
- Salida: gráfica de la función de distribución acumulada de los tamaños de los objetos

2.5.2.13. Simular una proxy-cache

- Descripción: Invoca a la aplicación de simulación de caches explicada en el [capítulo 5](#).
- Entrada: scripts de lanzamiento de la aplicación de simulación de caches situados en el directorio de trabajo de la aplicación
- Salida: hits y hits bytes resultantes

2.5.2.14. Salir

- Descripción: esta opción termina la ejecución de la aplicación.
- Entrada: ninguna
- Salida: ninguna

2.5.3. EJECUCIÓN POR LOTES

El estudiar las trazas por tipo de objeto puede resultar una tarea tediosa. Por ejemplo, para cada tipo que tenga una importante aportación al flujo de información (tales como html, jpg, gif,...), podría resultar útil calcular la media de los tamaños, el porcentaje de frecuencia con la que se solicita, y porcentaje del tamaño con respecto al total. Sería útil también, para una clase de objetos dada, calcular la media, moda, y desviación típica de los tamaños, de las tasas de transferencia, de los tiempos de respuesta,... Para realizar estas tareas de manera automática existe dos aplicaciones basadas en la aplicación principal (y que por tanto también corren bajo matlab): ExtBatch y ExtMainBatch. La primera de ellas realiza todos los cálculos anteriormente mencionados y para una clase de objetos especificado. En el caso de la segunda los cálculos se realizan solo para aquellos tipos que más aporten al tráfico.

CAPÍTULO 3: Estudio estadístico

El contenido de este tema abarca gran parte del objetivo que persigue este proyecto. Haciendo uso de gran parte de las herramientas anteriormente expuestas, se pretende descubrir algunas de las características principales del tráfico Web. Porcentajes de peticiones, medias, medianas, desviaciones típicas de los tamaños, de los tiempos de descarga, y de las tasas de transferencia para las distintas clases y tipos de objetos; son solo algunas de estas características.

3.1. ESTUDIO DE TRAZAS EN GENERAL

Se trata del estudio en general de cada uno de los archivos de traza, es decir, sin diferenciar tipos de objetos. Esto nos permite hacernos una idea de los archivos de traza que se estudian en profundidad en posteriores apartados. Para cada uno de los ficheros de trazas se obtienen:

- Porcentajes de frecuencias de los códigos y métodos HTTP.
- Medias, desviaciones típicas, máximos, y mínimos de los tamaños de los objetos y de los tiempos de descarga
- Número de trazas que componen los archivos.

A lo largo de este capítulo mostraremos todos los datos obtenidos como resultado de examinar todas las trazas (de la semana del 22/03/04 al 28/03/04) de ambos servidores (PA y RTP), haciendo uso de la opción “Estudiar una traza en general” de la aplicación principal, expuesta en el [apartado 2.5.2.7](#). Además, para cada uno de estos servidores, elaboraremos gráficas que resuman y comentaremos los datos de toda la semana.

3.1.1. SERVIDOR RTP

En las siguientes tablas se detallan datos estadísticos referentes a las peticiones del servidor proxy RTP. En cada una de estas tablas, una por cada día del estudio, se muestran:

- Porcentaje de aparición para cada uno de los códigos http.

- Media, Mediana, Desviación típica, máximos, y mínimos; de los tamaños y de los tiempos de descarga.
- Porcentaje de aparición de cada uno de los métodos (get, post, y head).

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
45,87%	38,64%	5,2%	10,28%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
15,2195	387,4251	0,1582	279770	13049000
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
1,5787	40,272	0	10208	1353600
METODOS				
GET	POST	HEAD	Total	
99,64%	0,04%	0,02%	1951158	

Tabla 3.1. Estudio general para las trazas del 22/03/04 del servidor RTP.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
47,55%	39,74%	2,8%	9,9%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
14,85	388,39	0,16	258810	2072900
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
1,72	42,88	0,00	10800	1562100
METODOS				
GET	POST	HEAD	Total	
99,96	0,03%	0,01%	1991950	

Tabla 3.2. Estudio general para las trazas del 23/03/04 del servidor RTP.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
45,67%	38,77%	5,94%	9,62%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
14,16	253,25	0,16	151090	13011000
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
1,55	56,31	0,00	10800	1427700
METODOS				
GET	POST	HEAD	Total	
99,62%	0,4%	0,2%	2096676	

Tabla 3.3. Estudio general para las trazas del 24/03/04 del servidor RTP.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
46,2%	39,11%	4,43%	10,25%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
14,70	308,97	0,16	244890	13047000
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
1,92	70,79	0,00	10800	1700000
METODOS				
GET	POST	HEAD	Total	
99,63%	0,4%	0,1%	2005748	

Tabla 3.4. Estudio general para las trazas del 25/03/04 del servidor RTP.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
43,61%	39,04%	6,5%	10,84%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
15,87	453,72	0,16	270390	12881000
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
1,93	62,67	0,00	10800	1564900
METODOS				
GET	POST	HEAD	Total	
99,53%	0,4%	0,2%	1953758	

Tabla 3.5. Estudio general para las trazas del 26/03/04 del servidor RTP.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
45,99%	33,74%	8,12%	12,15%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
15,57	151,52	0,16	32535	8154600
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
2,26	41,98	0,00	10799	1181900
METODOS				
GET	POST	HEAD	Total	
99,28%	0,6%	0,1%	1194143	

Tabla 3.6. Estudio general para las trazas del 27/03/04 del servidor RTP.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
39,16%	37,95%	8,6%	14,3%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
16,8067	128,0685	0,1602	18758	7902300
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
3,653	48,9479	0	5075	1717600
METODOS				
GET	POST	HEAD	Total	
99,4%	0,5%	0,1%	1265907	

Tabla 3.7. Estudio general para las trazas del 28/03/04 del servidor RTP.

3.1.1.1. Resumen

Para una mejor visualización de los datos anteriores se elaboran sus medias ponderadas (donde el peso es aportado por el número de trazas que componen un determinado día). De esta forma se observa que los códigos 2xx de operación exitosa y 3xx de redirección forman más del ochenta por ciento de las respuestas, tal y como muestra la figura 3.1. Así mismo en la figura 3.2 se observa como el método ‘GET’ es el más usado, ocupando más del noventa y nueve por ciento de las peticiones.

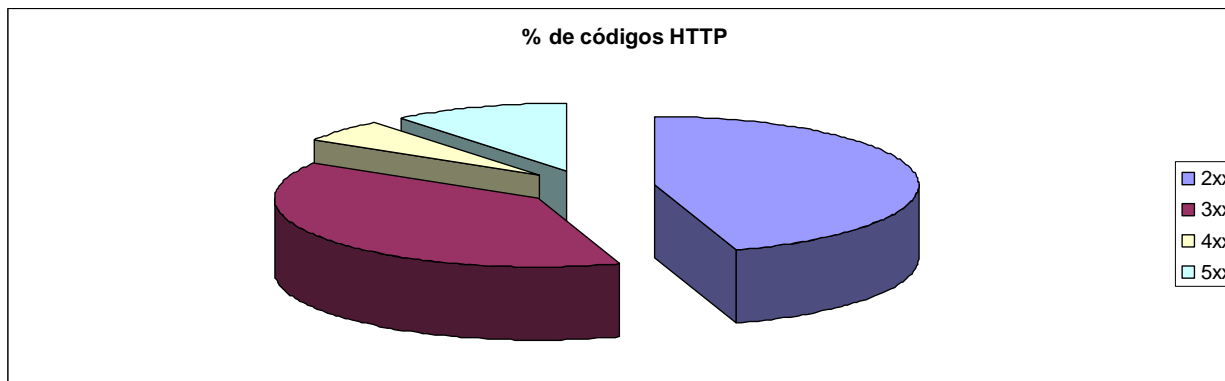


Figura 3.1. Gráfico de frecuencias de los códigos HTTP

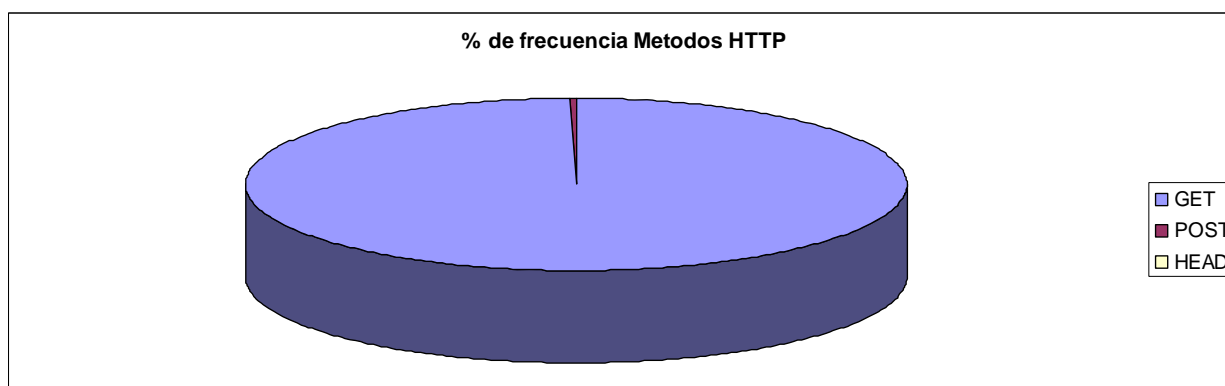


Figura 3.2. Gráfico de frecuencias de los métodos HTTP

En la tabla 3.8 encontramos el resto de resultados de resumen, donde las medias y desviaciones típicas de los tamaños y de los tiempos de descarga se han obtenido a partir de la media ponderada de toda la semana y donde el peso es aportado por el número de trazas que componen cada día. Para el caso de los mínimos y máximos, estos son los máximos y mínimos absolutos de toda la semana.

Tamaños (en KB)				Tiempos de descarga (sg)			
Media	Desviación Típica	Máximo	Mínimo	Media	Desviación Típica	Máximo	Mínimo
15,20	313,81	279770,00	0,16	1,98	52,86	10800,00	0,00

Tabla 3.8. Estadísticos de resumen para todas las trazas del servidor RTP

3.1.2. SERVIDOR PA

Al igual que se hizo en el apartado anterior para el caso del servidor proxy RTP, se detallan a continuación los datos estadísticos para las trazas del servidor proxy PA.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
21,15%	9,1%	0,6%	69,16%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
21,1715	140,3184	0,208	18441	2072900
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
5,4045	89,1302	0	6969	529150
METODOS				
GET	POST	HEAD	Total	
99,93%	0,06%	0,01%	476905	

Tabla 3.9. Estudio general para las trazas del 22/03/04 del servidor PA.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
20%	8,6%	1,4%	70,99%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
20,59	200,50	0,20	39941	1936600
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
4,38	63,00	0,00	4404,70	411980
METODOS				
GET	POST	HEAD	Total	
99,92%	0,05%	0,01%	506292	

Tabla 3.10. Estudio general para las trazas del 23/03/04 del servidor PA.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
29,19%	12%	0,92%	57,87%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
25,91	158,07	0,20	16368	1864800
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
4,96	73,18	0,00	10800	356860
METODOS				
GET	POST	HEAD	Total	
99,77%	0,01%	0%	254736	

Tabla 3.11. Estudio general para las trazas del 24/03/04 del servidor PA.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
39,22%	13,55%	1,26%	45,95%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
21,52	188,67	0,20	26616	1555400
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
5,64	73,72	0,00	5565	407800
METODOS				
GET	POST	HEAD	Total	
99,54%	0,3%	0%	188542	

Tabla 3.12. Estudio general para las trazas del 25/03/04 del servidor PA.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
40,04%	12,98%	1,42%	45,56%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
26,90	413,93	0,20	45907	1984800
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
8,22	170,34	0,00	10800	606230
METODOS				
GET	POST	HEAD	Total	
99,84%	0,15%	0,01%	191907	

Tabla 3.13. Estudio general para las trazas del 26/03/04 del servidor PA.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
40,16%	11,96%	1,2%	46,69%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
27,19	261,26	0,20	25713	1581500
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
7,66	160,04	0,00	10799	445730
METODOS				
GET	POST	HEAD	Total	
99,79%	0,21%	0%	148682	

Tabla 3.14. Estudio general para las trazas del 27/03/04 del servidor PA.

FRECUENCIAS CODIGOS HTTP				
2XX	3XX	4XX	5XX	
36,80%	22%	1,1%	40,1%	
TAMAÑOS (KB)				
Media	Desv. Típica	Mínimo	Máximo	Total
23,9772	202,5069	0,2051	19406	1458900
TIEMPOS DE DESCARGA (sg)				
Media	Desv. Típica	Mínimo	Máximo	Total
3,8704	63,1969	0	4918	235500
METODOS				
GET	POST	HEAD	Total	
99,84%	0,14%	0,01%	171143	

Tabla 3.15. Estudio general para las trazas del 28/03/04 del servidor PA.

3.1.2.1. Resumen

A diferencia del servidor proxy RTP, en el PA se observa que el código HTTP 500 (error por parte del servidor) forma la mayoría de las respuestas, tal y como se observa en la figura 3.3. No así, el método HTTP *Get* sigue formando más del noventa y nueve por ciento de las peticiones (figura 3.4), tal y como ocurría en el caso RTP.

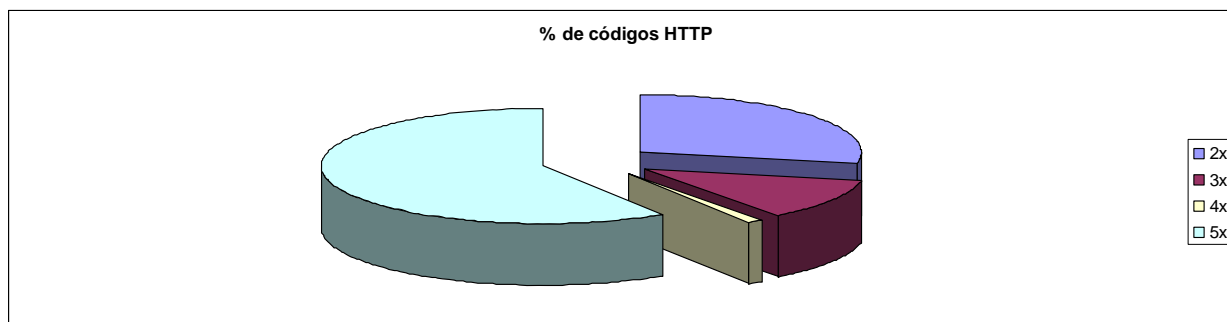


Figura 3.3. Gráfico de frecuencias de los códigos HTTP

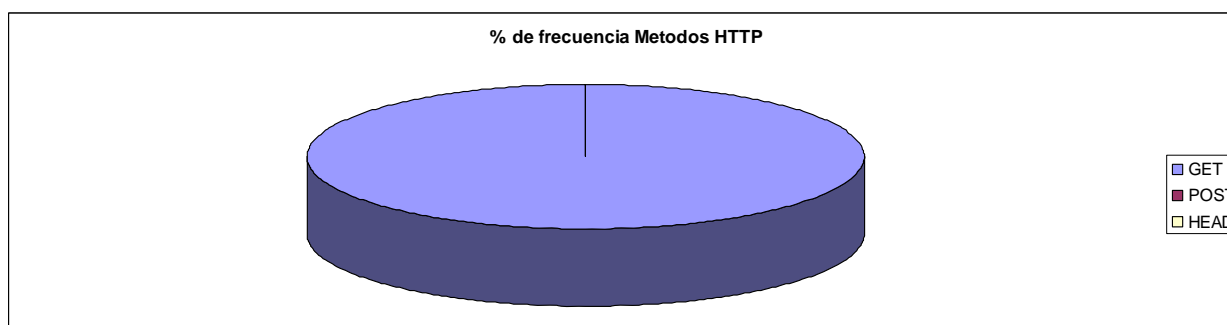


Figura 3.4. Gráfico de frecuencias de los métodos HTTP

de los tamaños (en KB)				de los tiempos de descarga (sg)			
Media	Desviación Típica	Máximo	Mínimo	Media	Desviación Típica	Máximo	Mínimo
22,95	204,93	45907,00	0,20	5,42	89,90	10800,00	0,00

Tabla 3.16. Estadísticos de resumen para el servidor PA

Las siguientes diferencias las encontramos en la tabla 3.16, donde el tamaño medio de los objetos es mayor que para el caso RTP, aunque la desviación típica es menor. Las diferencias también se extienden a los tiempos de descarga llegando a ser estos casi el doble mas lentos que para el caso RTP.

Es lógico pensar que si aumenta el tamaño medio de los objetos aumente también en consecuencia los tiempos de descarga, sin embargo la diferencia entre tamaños es pequeña y los tiempos de descarga del servidor PA llegan a duplicar el caso RTP.

Todo esto hace pensar en que el servidor proxy PA es mas pobre y con menor ancho de banda que el RTP o que sus usuarios tienen una conexión más lenta.

3.2. ESTUDIO DE TRAZAS SEGÚN CLASES Y TIPOS DE OBJETOS

Con el objetivo de resumir y poder comentar todos los resultados obtenidos, se han elaborado graficas basadas en las medias ponderadas de toda la semana estudiada, usando como peso el número de trazas que componen cada día. Además los datos obtenidos de ambos servidores proxy se han representado conjuntamente para poder compararlo con mayor facilidad.

3.2.1. ESTUDIO POR CLASE DE OBJETO

Se denominan clases de objetos a las distintas agrupaciones de tipos de objetos que tengan la misma funcionalidad. De esta forma, se agrupan objetos de tipo jpg, gif, bmp, ... en la clase imagen; avi, mpg, mov, ... en la clase video, ... En el apéndice de [clasificación de extensiones](#) se establecen las distintas clases de objeto que se ha utilizado para el proyecto, así como los tipos que forman cada una de ellas.

A continuación se exponen los datos obtenidos en el análisis de los servidores proxy RTP y PA a lo largo de la semana del 22 de Marzo de 2004 hasta 28 de Marzo de 2004. Estos resultados se han obtenido haciendo uso de las opciones ‘Porcentaje de frecuencias’, ‘Porcentaje de bytes’, y ‘Media, mediana, y desviación típica’, de la aplicación principal.

3.2.1.1. Servidor RTP

En este apartado se exponen las tablas de estadísticos como resultado del estudio de las trazas del servidor proxy RTP. Dichas tablas se encuentran ordenadas por fecha.

IMÁGENES	(%)		
Frecuencia	66,94		
Bytes	27,26		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,86	144,49	0,65
Mediana	2,45	33,41	0,08
Desv. Típica	30,69	469,87	11,74
Variabilidad	4,47	3,25	18,14
Mínimo	0,16	0,00	0,00
Máximo	9446,00	13525,00	1924,40
COMPRIMIDOS	(%)		
Frecuencia	0,71		
Bytes	2,69		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	34,30	354,63	4,35
Mediana	1,37	273,63	0,01
Desv. Típica	1569,60	307,86	66,58
Variabilidad	45,77	0,87	15,32
Mínimo	0,34	0,01	0,00
Máximo	155160,00	4554,00	2569,20
APLICACIÓN	(%)		
Frecuencia	3,57		
Bytes	11,77		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	31,95	300,76	3,36
Mediana	4,38	167,73	0,03
Desv. Típica	1321,20	418,93	86,84
Variabilidad	41,36	1,39	25,87
Mínimo	0,18	0,00	0,00
Máximo	279770,00	7894,50	10280,00
VIDEO	(%)		
Frecuencia	0,15		
Bytes	8,24		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	443,15	77,71	62,86
Mediana	86,70	7,56	5,56
Desv. Típica	1106,50	357,10	250,20
Variabilidad	2,50	4,60	3,98
Mínimo	0,27	0,01	0,00
Máximo	18934,00	6203,50	5979,20

SONIDO	(%)		
Frecuencia	0,17		
Bytes	1,36		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	84,54	592,67	50,71
Mediana	5,26	37,99	0,10
Desv. Típica	339,35	1620,30	469,77
Variabilidad	4,01	2,73	9,26
Mínimo	0,34	0,01	0,00
Máximo	6666,40	12551,00	9251,20
HTML	(%)		
Frecuencia	10,99		
Bytes	12,36		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	14,81	64,74	1,44
Mediana	4,94	19,95	0,20
Desv. Típica	22,74	231,84	18,18
Variabilidad	1,54	3,58	12,64
Mínimo	0,16	0,00	0,00
Máximo	2048,60	10010,00	1821,30
HTML DINAM.	(%)		
Frecuencia	4,34		
Bytes	3,75		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	11,36	32,70	1,55
Mediana	2,29	13,20	0,19
Desv. Típica	68,74	66,64	33,80
Variabilidad	6,05	2,04	21,75
Mínimo	0,18	0,00	0,00
Máximo	13500,00	5994,10	5734,20
TEXTO	(%)		
Frecuencia	1,12		
Bytes	3,42		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	30,06	182,95	2,90
Mediana	2,39	15,63	0,17
Desv. Típica	181,68	629,63	34,74
Variabilidad	6,04	3,44	12,00
Mínimo	0,16	0,00	0,00
Máximo	10441,00	10108,00	1810,10

HTTP	(%)		
Frecuencia	0,02		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,45	7,59	36,85
Mediana	0,33	7,95	0,06
Desv. Típica	0,28	7,06	126,45
Variabilidad	0,63	0,93	3,43
Mínimo	0,29	0,00	0,03
Máximo	2,48	43,09	703,36
OTROS	(%)		
Frecuencia	1,18		
Bytes	1,53		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	17,65	1177,70	2,08
Mediana	8,37	836,72	0,01
Desv. Típica	95,40	1272,50	30,22
Variabilidad	5,41	1,08	14,54
Mínimo	0,26	0,01	0,00
Máximo	4659,90	8396,50	2027,60
SCRIPT	(%)		
Frecuencia	3,81		
Bytes	1,16		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,80	182,82	0,59
Mediana	2,93	42,23	0,08
Desv. Típica	18,11	489,35	13,99
Variabilidad	2,32	2,68	23,60
Mínimo	0,25	0,00	0,00
Máximo	349,57	10805,00	1800,00
DOCUMENTOS	(%)		
Frecuencia	0,11		
Bytes	0,43		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	32,41	38,85	6,54
Mediana	8,26	19,20	0,30
Desv. Típica	191,55	88,03	75,44
Variabilidad	5,91	2,27	11,54
Mínimo	0,25	0,00	0,01
Máximo	5696,90	3219,90	1822,90

IRCACHE	(%)		
Frecuencia	0,44		
Bytes	18,31		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	1051,70	127,04	52,73
Mediana	752,04	84,94	6,20
Desv. Típica	933,59	127,80	143,37
Variabilidad	0,89	1,01	2,72
Mínimo	2,19	0,13	0,09
Máximo	2019,70	654,48	2953,50
PRESENTACION	(%)		
Frecuencia	2,28		
Bytes	0,51		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	5,48	149,69	0,99
Mediana	3,12	33,94	0,10
Desv. Típica	7,76	453,50	23,13
Variabilidad	1,42	3,03	23,46
Mínimo	0,25	0,00	0,00
Máximo	121,27	8247,10	1799,60
DESCONOCIDOS	(%)		
Frecuencia	4,16		
Bytes	7,22		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	26,99	81,86	9,50
Mediana	1,87	19,16	0,07
Desv. Típica	443,76	181,51	160,87
Coef. Var. Pearson	16,44	2,22	16,94
Mínimo	0,14	0,00	0,00
Máximo	37638,00	6312,20	10800,00

Tabla 3.17: Estudio por clase de objeto de las trazas del 22/03/04

IMÁGENES	(%)		
Frecuencia	67,64		
Bytes	28,31		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,98	150,23	0,60
Mediana	2,43	34,76	0,07
Desv. Típica	30,38	467,99	12,21
Coef. var. de Pearson	4,35	3,12	20,41
Mínimo	0,20	48770,00	0,00
Máximo	4074,80	15422,00	2770,30
COMPRIMIDOS	(%)		
Frecuencia	0,95		
Bytes	2,89		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	25,56	388,74	4,72
Mediana	1,37	304,30	0,01
Desv. Típica	1648,30	328,89	110,19
Coef. var. de Pearson	64,49	0,85	23,36
Mínimo	0,34	0,01	0,00
Máximo	210010,00	6991,20	5700,10
APLICACIÓN	(%)		
Frecuencia	5,16		
Bytes	10,78		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	38,16	246,32	3,83
Mediana	3,36	102,54	0,04
Desv. Típica	1342,60	402,49	82,09
Coef. var. de Pearson	35,18	1,63	21,41
Mínimo	0,19	0,00	0,00
Máximo	258810,00	7896,50	10800,00
VIDEO	(%)		
Frecuencia	0,28		
Bytes	8,17		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	234,26	163,30	48,96
Mediana	42,31	168,40	0,22
Desv. Típica	1444,00	212,16	256,17
Coef. var. de Pearson	6,16	1,30	5,23
Mínimo	0,27	0,00	0,00
Máximo	49122,00	6227,20	8117,40

SONIDO	(%)		
Frecuencia	0,22		
Bytes	2,17		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	111,61	648,67	66,98
Mediana	5,26	60,20	0,10
Desv. Típica	471,90	1528,70	380,15
Coef. var. de Pearson	4,23	2,36	5,68
Mínimo	0,41	0,00	0,00
Máximo	6052,30	9329,60	8374,85
HTML	(%)		
Frecuencia	8,62		
Bytes	12,04		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	15,11	64,01	1,70
Mediana	5,19	20,55	0,21
Desv. Típica	22,92	220,70	29,09
Coef. var. de Pearson	1,52	3,45	17,14
Mínimo	0,16	0,00	0,00
Máximo	1121,00	9383,30	2028,90
HTML DINAM.	(%)		
Frecuencia	3,64		
Bytes	3,53		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	8,62	26,13	1,45
Mediana	1,00	7,48	0,13
Desv. Típica	27,42	60,28	29,62
Coef. var. de Pearson	3,18	2,31	20,43
Mínimo	0,18	0,00	0,00
Máximo	2246,10	4454,10	1849,40
TEXTO	(%)		
Frecuencia	1,09		
Bytes	2,89		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	27,45	157,68	3,63
Mediana	2,38	15,29	0,15
Desv. Típica	162,52	564,70	47,81
Coef. var. de Pearson	5,92	3,58	13,17
Mínimo	0,16	0,00	0,00
Máximo	6020,70	10109,00	1811,30

HTTP	(%)		
Frecuencia	0,01		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,39	5,78	29,01
Mediana	0,29	6,15	0,05
Desv. Típica	0,33	4,78	100,65
Coef. var. de Pearson	0,85	0,83	3,47
Mínimo	0,29	0,00	0,03
Máximo	2,48	16,09	608,73
OTROS	(%)		
Frecuencia	1,28		
Bytes	1,83		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	15,03	1350,50	1,92
Mediana	8,37	1046,00	0,01
Desv. Típica	110,34	1292,40	61,07
Coef. var. de Pearson	7,34	0,96	31,85
Mínimo	0,24	0,01	0,00
Máximo	6118,40	8397,50	6430,50
SCRIPT	(%)		
Frecuencia	3,98		
Bytes	1,12		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,54	194,46	0,56
Mediana	3,01	44,53	0,08
Desv. Típica	16,73	525,32	17,96
Coef. var. de Pearson	2,22	2,70	32,04
Mínimo	0,25	0,00	0,00
Máximo	230,99	8052,70	1854,70
DOCUMENTOS	(%)		
Frecuencia	0,10		
Bytes	0,28		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	24,51	37,53	1,93
Mediana	8,47	15,36	0,26
Desv. Típica	132,25	49,79	10,28
Coef. var. de Pearson	5,40	1,33	5,32
Mínimo	0,26	0,01	0,01
Máximo	4378,40	593,75	291,08

IRCACHE	(%)		
Frecuencia	0,87		
Bytes	17,65		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	1044,50	126,20	51,29
Mediana	606,35	83,10	6,50
Desv. Típica	939,45	128,50	154,09
Coef. var. de Pearson	0,90	1,02	3,00
Mínimo	2,76	0,06	0,17
Máximo	2019,70	755,32	3121,60
PRESENTACION	(%)		
Frecuencia	2,38		
Bytes	0,55		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	5,70	147,06	1,33
Mediana	3,24	38,38	0,10
Desv. Típica	7,49	438,38	37,35
Coef. var. de Pearson	1,31	2,98	28,00
Mínimo	0,22	0,00	0,00
Máximo	119,84	8474,90	1800,10
DESCONOCIDOS	(%)		
Frecuencia	3,73		
Bytes	7,68		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	28,14	79,33	8,77
Mediana	1,87	20,06	0,08
Desv. Típica	522,69	205,29	148,89
Coef. var. de Pearson	18,57	2,59	16,98
Mínimo	0,14	0,00	0,00
Máximo	62180,00	7729,00	10800,00

Tabla 3.18: Estudio por clase de objeto de las trazas del 23/03/04

IMÁGENES	(%)		
Frecuencia	67,35		
Bytes	30,67		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,96	144,50	0,63
Mediana	2,54	35,81	0,07
Desv. Típica	32,73	440,03	12,57
Coef. var. de Pearson	4,70	3,05	19,80
Mínimo	0,20	0,00	0,00
Máximo	5898,90	15439,00	2209,50
COMPRIMIDOS	(%)		
Frecuencia	0,61		
Bytes	3,96		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	71,68	370,65	15,42
Mediana	1,37	274,22	0,01
Desv. Típica	2183,00	328,35	292,96
Coef. var. de Pearson	30,46	0,89	19,00
Mínimo	0,34	0,08	0,00
Máximo	151090,00	3614,00	10799,00
APLICACIÓN	(%)		
Frecuencia	4,75		
Bytes	9,13		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	34,44	223,93	5,08
Mediana	2,57	100,10	0,04
Desv. Típica	400,36	414,08	121,49
Coef. var. de Pearson	11,63	1,85	23,93
Mínimo	0,18	0,01	0,00
Máximo	31796,00	7895,50	10800,00
VIDEO	(%)		
Frecuencia	0,14		
Bytes	4,21		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	323,12	64,44	62,40
Mediana	35,35	7,64	0,83
Desv. Típica	970,19	392,20	205,08
Coef. var. de Pearson	3,00	6,09	3,29
Mínimo	0,26	0,00	0,00
Máximo	15771,00	9762,20	2141,90

SONIDO	(%)		
Frecuencia	0,21		
Bytes	2,30		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	132,66	425,22	33,04
Mediana	6,17	39,57	0,10
Desv. Típica	1388,30	1370,50	317,70
Coef. var. de Pearson	10,47	3,22	9,62
Mínimo	0,34	0,06	0,00
Máximo	59969,00	1173,10	9122,50
HTML	(%)		
Frecuencia	11,21		
Bytes	12,71		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	14,85	59,83	1,64
Mediana	5,21	21,48	0,21
Desv. Típica	23,12	182,45	19,67
Coef. var. de Pearson	1,56	3,05	12,03
Mínimo	0,16	0,00	0,00
Máximo	2049,50	7374,00	1894,50
HTML DINAM.	(%)		
Frecuencia	3,22		
Bytes	3,37		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	8,57	28,54	1,53
Mediana	1,38	10,20	0,15
Desv. Típica	20,29	61,93	26,14
Coef. var. de Pearson	2,37	2,17	17,11
Mínimo	0,16	0,00	0,00
Máximo	1506,80	4442,40	1802,60
TEXTO	(%)		
Frecuencia	1,15		
Bytes	4,07		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	34,52	110,27	3,64
Mediana	2,39	29,09	0,08
Desv. Típica	224,87	437,18	43,99
Coef. var. de Pearson	6,51	3,96	12,07
Mínimo	0,16	0,00	0,00
Máximo	1111,50	8249,70	1875,60

HTTP	(%)		
Frecuencia	0,02		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,35	8,10	16,73
Mediana	0,29	9,51	0,03
Desv. Típica	0,27	4,43	79,41
Coef. var. de Pearson	0,75	0,55	4,75
Mínimo	0,29	0,00	0,03
Máximo	2,62	46,01	594,37
OTROS	(%)		
Frecuencia	1,07		
Bytes	1,40		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	16,86	992,75	1,62
Mediana	8,37	464,95	0,01
Desv. Típica	463,19	1221,90	36,35
Coef. var. de Pearson	27,47	1,23	22,46
Mínimo	0,29	0,01	0,00
Máximo	4883,20	8394,50	1900,30
SCRIPT	(%)		
Frecuencia	3,67		
Bytes	1,23		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,98	194,53	0,53
Mediana	3,17	47,41	0,07
Desv. Típica	18,21	514,43	13,45
Coef. var. de Pearson	2,28	2,64	25,42
Mínimo	0,25	0,00	0,00
Máximo	388,24	12828,00	1800,90
DOCUMENTOS	(%)		
Frecuencia	0,09		
Bytes	0,36		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	32,95	42,26	7,64
Mediana	12,95	17,39	0,35
Desv. Típica	159,00	59,51	142,17
Coef. var. de Pearson	4,83	1,41	18,60
Mínimo	0,26	0,00	0,01
Máximo	3338,40	575,32	5429,40

IRCACHE	(%)		
Frecuencia	0,82		
Bytes	17,82		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	1061,90	128,93	47,66
Mediana	696,04	86,42	6,58
Desv. Típica	936,04	129,55	120,21
Coef. var. de Pearson	0,88	1,00	2,52
Mínimo	0,82	0,10	0,02
Máximo	2019,70	652,58	844,18
PRESENTACION	(%)		
Frecuencia	2,37		
Bytes	0,52		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	5,54	150,03	1,41
Mediana	3,29	37,28	0,10
Desv. Típica	7,31	432,12	37,01
Coef. var. de Pearson	1,32	2,88	26,16
Mínimo	0,23	0,00	0,00
Máximo	122,79	9877,90	1806,00
DESCONOCIDOS	(%)		
Frecuencia	3,27		
Bytes	8,19		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	30,29	52,67	8,04
Mediana	1,36	13,71	0,10
Desv. Típica	586,37	141,67	129,17
Coef. var. de Pearson	19,36	2,69	16,06
Mínimo	0,14	0,00	0,00
Máximo	44065,00	6208,00	8043,70

Tabla 3.19: Estudio por clase de objeto de las trazas del 24/03/04

IMÁGENES	(%)		
Frecuencia	65,49		
Bytes	30,34		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,48	144,91	0,63
Mediana	2,60	37,09	0,07
Desv. Típica	36,65	436,16	12,97
Variabilidad	4,90	3,01	20,63
Mínimo	0,16	0,00	0,00
Máximo	8302,90	13545,00	2592,70
COMPRESION	(%)		
Frecuencia	0,68		
Bytes	3,43		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	52,60	373,54	47,70
Mediana	1,37	274,02	0,01
Desv. Típica	1075,20	319,97	591,87
Variabilidad	20,44	0,86	12,41
Mínimo	0,34	0,09	0,00
Máximo	73440,00	4518,80	10800,00
APLICACIÓN	(%)		
Frecuencia	7,92		
Bytes	9,27		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	19,28	261,48	3,01
Mediana	3,27	146,48	0,03
Desv. Típica	269,83	385,56	94,23
Variabilidad	14,00	1,47	31,29
Mínimo	0,19	0,00	0,00
Máximo	30942,00	7895,50	10800,00
VIDEO	(%)		
Frecuencia	0,11		
Bytes	6,47		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	477,59	32,42	64,54
Mediana	36,77	7,45	1,83
Desv. Típica	5658,80	80,94	213,37
Variabilidad	11,85	2,50	3,31
Mínimo	0,26	0,01	0,00
Máximo	244890,00	914,51	2955,90

SONIDO	(%)		
Frecuencia	0,20		
Bytes	2,49		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	161,55	575,14	47,69
Mediana	6,65	37,91	0,10
Desv. Típica	777,89	1516,40	313,09
Variabilidad	4,82	2,64	6,56
Mínimo	0,26	0,08	0,00
Máximo	12006,00	10258,00	6754,10
HTML	(%)		
Frecuencia	8,83		
Bytes	12,55		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	15,49	70,06	1,31
Mediana	5,24	22,24	0,21
Desv. Típica	26,28	284,92	21,97
Variabilidad	1,70	4,07	16,71
Mínimo	0,16	0,00	0,00
Máximo	3082,50	9920,40	2752,40
HTML DINAM.	(%)		
Frecuencia	3,39		
Bytes	3,87		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	9,93	28,68	2,17
Mediana	1,37	9,66	0,14
Desv. Típica	34,64	64,31	33,46
Variabilidad	3,49	2,24	15,43
Mínimo	0,18	0,00	0,00
Máximo	2309,70	4755,90	1800,80
TEXTO	(%)		
Frecuencia	1,20		
Bytes	3,10		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	29,33	117,52	2,75
Mediana	1,90	13,94	0,15
Desv. Típica	150,63	500,16	29,60
Variabilidad	5,14	4,26	10,76
Mínimo	0,18	0,00	0,00
Máximo	5832,70	7907,00	1561,50

HTTP	(%)		
Frecuencia	0,02		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,37	8,80	4,29
Mediana	0,29	9,66	0,03
Desv. Típica	0,34	5,18	41,53
Variabilidad	0,93	0,59	9,67
Mínimo	0,28	0,00	0,03
Máximo	5,19	42,55	604,20
OTROS	(%)		
Frecuencia	1,24		
Bytes	1,27		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	14,15	1183,30	1,73
Mediana	8,37	836,43	0,01
Desv. Típica	126,46	1282,60	38,25
Variabilidad	8,94	1,08	22,08
Mínimo	0,24	0,01	0,00
Máximo	5447,60	5291,00	2118,70
SCRIPT	(%)		
Frecuencia	3,65		
Bytes	1,12		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,37	198,39	0,37
Mediana	3,04	47,50	0,07
Desv. Típica	14,89	526,73	3,59
Variabilidad	2,02	2,66	9,79
Mínimo	0,25	0,01	0,00
Máximo	266,88	8048,80	185,17
DOCUMENTOS	(%)		
Frecuencia	0,10		
Bytes	0,25		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	21,08	39,84	2,66
Mediana	9,56	15,32	0,32
Desv. Típica	46,31	58,78	22,02
Variabilidad	2,20	1,48	8,27
Mínimo	0,25	0,00	0,00
Máximo	869,86	977,54	833,62

IRCACHE	(%)		
Frecuencia	1,40		
Bytes	17,97		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	1097,90	130,74	46,15
Mediana	2019,70	88,60	6,55
Desv. Típica	936,27	129,25	113,66
Variabilidad	0,85	0,99	2,46
Mínimo	5,43	0,13	0,19
Máximo	2019,70	650,48	684,49
PRESENTACION	(%)		
Frecuencia	2,23		
Bytes	0,47		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	5,37	127,78	1,63
Mediana	3,38	32,28	0,11
Desv. Típica	6,44	337,85	37,52
Variabilidad	1,20	2,64	23,08
Mínimo	0,26	0,00	0,00
Máximo	117,45	7439,50	1800,20
DESCONOCIDOS	(%)		
Frecuencia	3,54		
Bytes	7,39		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	26,85	58,63	8,92
Mediana	1,45	14,85	0,08
Desv. Típica	383,16	143,71	145,55
Variabilidad	14,27	2,45	16,32
Mínimo	0,14	0,00	0,00
Máximo	26834,00	7117,20	1079,90

Tabla 3.20: Estudio por clase de objeto de las trazas del 25/03/04

IMÁGENES	(%)		
Frecuencia	64,56		
Bytes	26,57		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,26	144,57	0,66
Mediana	2,53	36,51	0,07
Desv. Típica	50,64	448,28	20,31
Variabilidad	6,97	3,10	30,59
Mínimo	0,20	0,00	0,00
Máximo	26452,00	12521,00	6289,70
COMPRESION	(%)		
Frecuencia	0,65		
Bytes	3,34		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	52,24	371,77	32,55
Mediana	1,37	274,22	0,01
Desv. Típica	1013,00	347,91	456,65
Variabilidad	19,39	0,94	14,03
Mínimo	0,34	0,01	0,00
Máximo	73022,00	7445,10	10800,00
APLICACIÓN	(%)		
Frecuencia	6,07		
Bytes	15,27		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	46,38	206,39	4,64
Mediana	3,04	96,79	0,04
Desv. Típica	1790,20	357,68	124,71
Variabilidad	38,60	1,73	26,86
Mínimo	0,19	0,00	0,00
Máximo	270390,00	8096,70	10800,00
VIDEO	(%)		
Frecuencia	0,27		
Bytes	5,49		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	156,68	142,35	25,83
Mediana	32,01	142,55	0,16
Desv. Típica	611,53	153,65	147,56
Variabilidad	3,90	1,08	5,71
Mínimo	0,26	0,00	0,00
Máximo	13882,00	4822,00	2834,40

SONIDO	(%)		
Frecuencia	0,19		
Bytes	1,17		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	71,88	908,42	12,59
Mediana	5,36	75,01	0,01
Desv. Típica	339,58	1943,90	104,57
Variabilidad	4,72	2,14	8,30
Mínimo	0,36	0,01	0,00
Máximo	7814,40	11735,00	2264,20
HTML	(%)		
Frecuencia	11,69		
Bytes	12,46		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	15,35	69,44	1,53
Mediana	5,21	22,75	0,21
Desv. Típica	25,79	269,06	24,96
Variabilidad	1,68	3,87	16,29
Mínimo	0,16	0,00	0,00
Máximo	3258,80	9867,20	2227,80
HTML DINAM.	(%)		
Frecuencia	3,45		
Bytes	3,96		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	10,20	28,65	1,78
Mediana	1,76	10,64	0,20
Desv. Típica	27,46	68,50	27,18
Variabilidad	2,69	2,39	15,37
Mínimo	0,18	0,00	0,00
Máximo	2556,30	6228,50	1945,70
TEXTO	(%)		
Frecuencia	1,17		
Bytes	2,87		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	27,21	157,00	7,56
Mediana	0,97	10,04	0,22
Desv. Típica	170,28	710,64	43,43
Variabilidad	6,26	4,53	5,74
Mínimo	0,18	0,00	0,00
Máximo	8116,90	10527,00	1400,20

HTTP	(%)		
Frecuencia	0,02		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,35	7,96	39,35
Mediana	0,29	9,51	0,03
Desv. Típica	0,22	4,88	119,83
Variabilidad	0,64	0,61	3,05
Mínimo	0,28	0,00	0,03
Máximo	2,48	37,38	691,97
OTROS	(%)		
Frecuencia	1,34		
Bytes	1,13		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	11,06	1200,80	1,28
Mediana	8,37	911,62	0,01
Desv. Típica	84,69	1245,90	36,19
Variabilidad	7,66	1,04	28,20
Mínimo	0,25	0,02	0,00
Máximo	5447,60	4199,70	2657,70
SCRIPT	(%)		
Frecuencia	3,66		
Bytes	0,91		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,86	171,61	0,45
Mediana	2,81	41,32	0,07
Desv. Típica	14,71	452,36	13,74
Variabilidad	2,15	2,64	30,68
Mínimo	0,26	0,01	0,00
Máximo	388,24	14100,00	1823,80
DOCUMENTOS	(%)		
Frecuencia	0,10		
Bytes	0,39		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	30,96	45,81	5,48
Mediana	12,91	19,04	0,31
Desv. Típica	139,66	61,71	72,53
Variabilidad	4,51	1,35	13,23
Mínimo	0,25	0,00	0,02
Máximo	3529,50	460,77	2624,50

IRCACHE	(%)		
Frecuencia	0,85		
Bytes	17,31		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	1108,90	13,85	48,46
Mediana	2019,70	92,44	6,53
Desv. Típica	935,78	129,38	118,52
Variabilidad	0,84	0,97	2,45
Mínimo	5,44	0,12	0,04
Máximo	2019,70	651,95	722,71
PRESENTACION	(%)		
Frecuencia	2,21		
Bytes	0,42		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	5,39	149,59	1,27
Mediana	3,31	37,92	0,10
Desv. Típica	6,40	428,69	34,82
Variabilidad	1,19	2,87	27,49
Mínimo	0,22	0,00	0,00
Máximo	120,65	9835,00	1799,70
DESCONOCIDOS	(%)		
Frecuencia	3,77		
Bytes	8,71		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	30,27	88,16	8,48
Mediana	1,87	18,69	0,07
Desv. Típica	455,90	206,08	127,53
Variabilidad	15,06	2,34	15,03
Mínimo	0,14	0,00	0,00
Máximo	27056,00	10598,00	10799,00

Tabla 3.21: Estudio por clase de objeto de las trazas del 26/03/04

IMÁGENES	(%)		
Frecuencia	58,97		
Bytes	27,72		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,61	138,00	0,68
Mediana	2,77	38,06	0,07
Desv. Típica	34,29	438,88	16,16
Variabilidad	4,50	3,18	23,92
Mínimo	0,20	0,00	0,00
Máximo	4264,30	13243,00	2246,20
COMPRESION	(%)		
Frecuencia	0,94		
Bytes	2,32		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	22,33	330,55	7,11
Mediana	1,37	228,52	0,01
Desv. Típica	454,68	319,58	153,16
Variabilidad	20,34	0,97	21,54
Mínimo	0,34	0,01	0,00
Máximo	32535,00	6986,30	10799,00
APLICACIÓN	(%)		
Frecuencia	8,94		
Bytes	6,63		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	19,35	161,69	3,56
Mediana	24,59	92,84	0,03
Desv. Típica	248,26	253,29	87,69
Variabilidad	12,83	1,57	24,60
Mínimo	0,19	0,01	0,00
Máximo	18473,00	7895,50	10534,00
VIDEO	(%)		
Frecuencia	0,19		
Bytes	7,36		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	319,50	29,69	117,99
Mediana	70,80	3,22	53,73
Desv. Típica	1155,70	81,96	245,32
Variabilidad	3,62	2,76	2,08
Mínimo	0,27	0,00	0,00
Máximo	26311,00	997,72	2972,70

SONIDO	(%)		
Frecuencia	0,17		
Bytes	1,53		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	99,18	490,10	27,72
Mediana	7,22	38,54	0,10
Desv. Típica	316,34	1573,10	153,76
Variabilidad	3,19	3,21	5,55
Mínimo	0,34	0,00	0,00
Máximo	5829,70	11481,00	1800,20
HTML	(%)		
Frecuencia	12,72		
Bytes	11,38		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	13,53	57,42	1,55
Mediana	4,30	20,75	0,17
Desv. Típica	23,08	201,96	21,44
Variabilidad	1,71	3,52	13,83
Mínimo	0,16	0,00	0,00
Máximo	1245,10	11548,00	1959,90
HTML DINAM.	(%)		
Frecuencia	4,95		
Bytes	4,57		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,87	22,32	1,98
Mediana	1,10	9,43	0,18
Desv. Típica	17,90	43,42	32,34
Variabilidad	2,27	1,95	16,30
Mínimo	0,16	0,00	0,00
Máximo	802,54	14795,00	2018,20
TEXTO	(%)		
Frecuencia	1,70		
Bytes	1,62		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	9,27	74,28	10,18
Mediana	0,97	1,47	0,67
Desv. Típica	83,26	452,70	30,75
Variabilidad	8,98	6,09	3,02
Mínimo	0,18	0,00	0,00
Máximo	4289,50	8175,80	1096,00

HTTP	(%)		
Frecuencia	0,04		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,34	7,53	5,41
Mediana	0,28	8,01	0,04
Desv. Típica	0,20	4,09	46,59
Variabilidad	0,59	0,54	8,61
Mínimo	0,28	0,00	0,02
Máximo	2,44	22,58	510,91
OTROS	(%)		
Frecuencia	1,55		
Bytes	1,35		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	9,94	1010,60	1,58
Mediana	8,37	643,78	0,01
Desv. Típica	88,09	1173,30	48,88
Variabilidad	8,86	1,16	30,85
Mínimo	0,24	0,01	0,00
Máximo	5453,30	8399,40	3400,10
SCRIPT	(%)		
Frecuencia	2,88		
Bytes	0,69		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,59	164,29	0,38
Mediana	2,91	45,43	0,07
Desv. Típica	11,51	473,54	2,58
Variabilidad	1,75	2,88	6,78
Mínimo	0,32	0,01	0,00
Máximo	272,29	8404,80	89,15
DOCUMENTOS	(%)		
Frecuencia	0,10		
Bytes	0,26		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	21,70	39,52	5,95
Mediana	13,21	15,30	0,40
Desv. Típica	59,46	89,00	52,12
Variabilidad	2,74	2,25	8,76
Mínimo	0,26	0,01	0,00
Máximo	1546,90	2220,80	1492,90

IRCACHE	(%)		
Frecuencia	1,70		
Bytes	28,74		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	1240,00	141,19	62,85
Mediana	2019,70	93,57	7,85
Desv. Típica	928,36	133,36	192,63
Variabilidad	0,75	0,94	3,06
Mínimo	2,20	0,17	0,02
Máximo	2019,70	651,95	3370,80
PRESENTACION	(%)		
Frecuencia	1,60		
Bytes	0,33		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	5,05	123,90	1,47
Mediana	3,10	34,44	0,10
Desv. Típica	5,61	336,16	41,42
Variabilidad	1,11	2,71	28,10
Mínimo	0,29	0,00	0,00
Máximo	61,21	9968,80	1799,90
DESCONOCIDOS	(%)		
Frecuencia	3,55		
Bytes	5,49		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	18,62	94,22	9,33
Mediana	2,35	36,42	0,05
Desv. Típica	190,34	202,92	114,84
Variabilidad	10,22	2,15	11,56
Mínimo	0,14	0,00	0,00
Máximo	15933,00	8091,80	6894,60

Tabla 3.22: Estudio por clase de objeto de las trazas del 27/03/04

IMÁGENES	(%)		
Frecuencia	56,85		
Bytes	25,49		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,60	146,02	0,78
Mediana	2,71	39,73	0,07
Desv. Típica	31,56	499,92	14,55
Variabilidad	4,15	3,42	18,66
Mínimo	0,20	0,00	0,00
Máximo	4931,40	15833,00	2210,80
COMPRESION	(%)		
Frecuencia	0,55		
Bytes	1,11		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	31,93	341,62	9,81
Mediana	1,37	228,52	0,01
Desv. Típica	359,40	371,59	109,73
Variabilidad	11,26	1,09	11,19
Mínimo	0,33	0,20	0,00
Máximo	14649,00	5953,80	3076,10
APLICACIÓN	(%)		
Frecuencia	13,26		
Bytes	9,94		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	35,54	192,42	5,47
Mediana	4,40	95,79	0,05
Desv. Típica	228,88	332,47	91,26
Variabilidad	6,44	1,73	16,67
Mínimo	0,19	0,00	0,00
Máximo	18758,00	7894,50	5075,00
VIDEO	(%)		
Frecuencia	0,46		
Bytes	5,87		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	110,53	6,35	160,57
Mediana	51,00	0,41	119,51
Desv. Típica	344,13	46,78	242,06
Variabilidad	3,11	7,37	1,51
Mínimo	0,27	0,00	0,00
Máximo	15359,00	23154,00	3251,90

SONIDO	(%)		
Frecuencia	0,20		
Bytes	1,09		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	61,65	1488,50	98,63
Mediana	6,65	129,62	0,06
Desv. Típica	211,23	2659,70	385,37
Variabilidad	3,43	1,79	3,91
Mínimo	0,29	0,00	0,00
Máximo	3676,40	15350,00	2133,50
HTML	(%)		
Frecuencia	9,15		
Bytes	9,44		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	12,36	67,08	1,62
Mediana	3,47	21,63	0,15
Desv. Típica	24,12	350,01	21,34
Variabilidad	1,95	5,22	13,20
Mínimo	0,16	0,00	0,00
Máximo	2639,20	9918,90	1924,80
HTML DINAM.	(%)		
Frecuencia	4,49		
Bytes	5,37		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	9,87	24,43	2,13
Mediana	1,75	10,02	0,25
Desv. Típica	25,30	46,03	27,57
Variabilidad	2,56	1,88	12,97
Mínimo	0,18	0,00	0,00
Máximo	2856,30	2743,70	2103,30
TEXTO	(%)		
Frecuencia	0,91		
Bytes	1,34		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	12,98	130,01	8,86
Mediana	0,97	1,49	0,66
Desv. Típica	112,28	628,69	55,88
Variabilidad	8,65	4,84	6,31
Mínimo	0,18	0,00	0,00
Máximo	4681,50	7613,30	1881,30

HTTP	(%)		
Frecuencia	0,01		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,36	7,60	27,00
Mediana	0,29	9,20	0,03
Desv. Típica	0,29	4,01	97,39
Variabilidad	0,82	0,53	3,61
Mínimo	0,29	0,00	0,03
Máximo	2,66	19,17	491,35
OTROS	(%)		
Frecuencia	1,18		
Bytes	0,86		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	10,31	911,40	1,48
Mediana	8,37	522,83	0,01
Desv. Típica	68,27	1137,00	60,33
Variabilidad	6,62	1,25	40,84
Mínimo	0,26	0,01	0,00
Máximo	2738,00	8369,10	4833,70
SCRIPT	(%)		
Frecuencia	2,98		
Bytes	0,68		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,80	201,81	0,42
Mediana	3,00	45,40	0,07
Desv. Típica	12,91	631,94	4,09
Variabilidad	1,90	3,13	9,74
Mínimo	0,33	0,01	0,00
Máximo	388,23	10189,00	201,03
DOCUMENTOS	(%)		
Frecuencia	0,13		
Bytes	0,37		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	26,86	25,50	16,37
Mediana	17,24	12,40	0,63
Desv. Típica	97,35	55,90	172,74
Variabilidad	3,62	1,57	10,56
Mínimo	0,25	0,00	0,00
Máximo	2224,90	607,11	4329,80

IRCACHE	(%)		
Frecuencia	4,22		
Bytes	29,45		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	1232,50	140,49	56,66
Mediana	2019,70	93,61	7,14
Desv. Típica	930,32	132,17	140,74
Variabilidad	0,75	0,94	2,48
Mínimo	2,96	0,27	0,18
Máximo	2019,70	653,00	1627,90
PRESENTACION	(%)		
Frecuencia	1,50		
Bytes	0,35		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	5,61	156,79	1,55
Mediana	3,02	35,74	0,10
Desv. Típica	8,75	593,73	35,77
Variabilidad	1,56	3,79	23,05
Mínimo	0,31	0,00	0,00
Máximo	123,71	10486,00	18046,00
DESCONOCIDOS	(%)		
Frecuencia	4,11		
Bytes	8,65		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	29,20	79,70	5,26
Mediana	1,96	25,44	0,05
Desv. Típica	401,81	183,34	64,39
Variabilidad	13,76	2,30	12,25
Mínimo	0,14	0,00	0,00
Máximo	14649,00	7032,70	3280,80

Tabla 3.23: Estudio por clase de objeto de las trazas del 28/03/04

3.2.1.2. Servidor PA

IMÁGENES	(%)		
Frecuencia	32,45		
Bytes	27,64		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	9,59	44,35	1,19
Mediana	2,59	14,15	0,22
Desv. Típica	33,94	140,93	13,34
Coef. var. de Pearson	3,54	3,18	11,24
Mínimo	0,22	0,00	0,00
Máximo	1752,60	11611,00	2008,00
COMPRESION	(%)		
Frecuencia	0,08		
Bytes	2,95		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	403,47	113,07	222,31
Mediana	14,19	29,61	0,57
Desv. Típica	2179,30	228,07	539,66
Coef. var. de Pearson	5,40	2,02	2,43
Mínimo	0,07	0,00	0,00
Máximo	18441,00	1476,60	2393,60
APLICACIÓN	(%)		
Frecuencia	0,79		
Bytes	8,56		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	56,49	26,90	45,12
Mediana	0,76	19,20	0,05
Desv. Típica	365,10	38,80	403,04
Coef. var. de Pearson	6,46	1,44	8,93
Mínimo	0,29	0,01	0,01
Máximo	9958,90	468,81	6968,60

VIDEO	(%)		
Frecuencia	0,08		
Bytes	3,17		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	362,57	28,62	76,48
Mediana	140,86	6,91	10,89
Desv. Típica	532,57	48,87	157,39
Coef. var. de Pearson	1,47	1,71	2,06
Mínimo	0,34	0,07	0,02
Máximo	3660,90	311,66	1290,00
SONIDO	(%)		
Frecuencia	0,05		
Bytes	0,31		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	115,79	101,34	24,29
Mediana	14,17	27,60	0,30
Desv. Típica	440,62	404,94	88,03
Coef. var. de Pearson	3,81	4,00	3,62
Mínimo	0,68	0,41	0,01
Máximo	2786,70	3043,70	602,73
HTML	(%)		
Frecuencia	7,49		
Bytes	20,64		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	21,55	24,81	6,35
Mediana	15,63	10,90	0,87
Desv. Típica	23,03	74,31	29,37
Coef. var. de Pearson	1,07	3,00	4,63
Mínimo	0,21	0,00	0,00
Máximo	731,86	5874,80	1800,40
HTML DINAM.	(%)		
Frecuencia	54,24		
Bytes	2,80		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	10,45	25,07	2,37
Mediana	3,91	12,53	0,39
Desv. Típica	17,50	34,98	15,85
Coef. var. de Pearson	1,67	1,40	6,68
Mínimo	0,22	0,00	0,01
Máximo	495,71	581,12	947,23

TEXTO	(%)		
Frecuencia	0,37		
Bytes	1,67		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	34,06	70,82	6,05
Mediana	0,88	7,45	0,16
Desv. Típica	137,90	414,99	31,74
Coef. var. de Pearson	4,05	5,86	5,25
Mínimo	0,21	0,01	0,00
Máximo	2275,80	7582,30	582,97
HTTP	(%)		
Frecuencia	0,00		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,54	1,96	29,33
Mediana	0,40	2,23	0,19
Desv. Típica	0,35	1,65	85,72
Coef. var. de Pearson	0,65	0,84	2,92
Mínimo	0,30	0,00	0,15
Máximo	1,70	5,42	350,32
OTROS	(%)		
Frecuencia	0,50		
Bytes	0,63		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	9,60	98,99	2,41
Mediana	1,70	19,56	0,10
Desv. Típica	37,35	354,64	30,56
Coef. var. de Pearson	3,89	3,58	12,69
Mínimo	0,29	0,04	0,00
Máximo	599,41	4197,80	1076,00
SCRIPT	(%)		
Frecuencia	1,37		
Bytes	0,65		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,02	51,62	1,34
Mediana	2,86	16,55	0,22
Desv. Típica	10,58	117,97	10,74
Coef. var. de Pearson	1,51	2,29	7,99
Mínimo	0,30	0,01	0,00
Máximo	126,00	2963,80	256,30

DOCUMENTOS	(%)		
Frecuencia	0,07		
Bytes	0,37		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	39,47	32,48	12,60
Mediana	22,35	24,24	0,65
Desv. Típica	87,23	38,00	128,39
Coef. var. de Pearson	2,21	1,17	9,93
Mínimo	0,46	0,05	0,03
Máximo	816,37	307,33	1805,70
IRCACHE	(%)		
Frecuencia	0,21		
Bytes	28,15		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	606,98	91,73	104,84
Mediana	1025,90	67,90	7,80
Desv. Típica	455,74	96,19	386,87
Coef. var. de Pearson	0,75	1,05	3,69
Mínimo	40,10	0,05	0,29
Máximo	1025,90	926,72	4599,90
PRESENTACION	(%)		
Frecuencia	1,18		
Bytes	0,36		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	4,45	28,49	0,86
Mediana	2,57	10,10	0,27
Desv. Típica	5,70	56,01	3,52
Coef. var. de Pearson	1,28	1,97	4,10
Mínimo	0,30	0,03	0,00
Máximo	95,02	713,38	60,55
DESCONOCIDOS	(%)		
Frecuencia	1,12		
Bytes	2,12		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	15,09	38,52	5,33
Mediana	2,33	14,01	0,24
Desv. Típica	65,50	52,78	53,84
Coef. var. de Pearson	4,34	1,37	10,10
Mínimo	0,17	0,01	0,01
Máximo	1026,00	604,67	1611,70

Tabla 3.24: Estudio por clase de objeto de las trazas del 22/03/04

IMÁGENES	(%)		
Frecuencia	32,35		
Bytes	20,63		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,72	39,92	1,40
Mediana	2,68	13,99	0,23
Desv. Típica	36,18	108,68	24,64
Coef. var. de Pearson	4,69	2,72	17,66
Mínimo	0,20	0,00	0,00
Máximo	5327,20	6780,30	2012,70
COMPRESION	(%)		
Frecuencia	0,36		
Bytes	2,15		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	46,41	182,12	12,29
Mediana	1,40	69,21	0,02
Desv. Típica	408,70	311,82	97,00
Coef. var. de Pearson	8,81	1,71	7,89
Mínimo	0,38	0,13	0,00
Máximo	10812,00	3047,90	1812,70
APLICACIÓN	(%)		
Frecuencia	0,67		
Bytes	4,99		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	44,39	25,30	32,83
Mediana	0,76	16,99	0,06
Desv. Típica	442,60	377352,00	173,90
Coef. var. de Pearson	9,97	1,49	5,30
Mínimo	0,23	0,03	0,01
Máximo	16091,00	410,13	3848,70
VIDEO	(%)		
Frecuencia	0,08		
Bytes	6,10		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	607,10	41,68	70,60
Mediana	291,96	20,28	9,22
Desv. Típica	855,41	43,85	177,45
Coef. var. de Pearson	1,41	1,05	2,51
Mínimo	0,34	0,07	0,02
Máximo	5494,40	159,84	1469,00

SONIDO	(%)		
Frecuencia	0,12		
Bytes	3,81		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	1013,70	52,66	70,71
Mediana	36,85	22,35	2,38
Desv. Típica	4656,40	117,56	292,49
Coef. var. de Pearson	4,59	2,23	4,14
Mínimo	0,56	0,31	0,01
Máximo	39941,00	983,79	2200,40
HTML	(%)		
Frecuencia	7,14		
Bytes	13,15		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	18,31	28,23	3,34
Mediana	7,55	14,96	0,51
Desv. Típica	24,75	60,97	45,09
Coef. var. de Pearson	1,35	2,16	13,52
Mínimo	0,21	0,00	0,00
Máximo	784,69	3087,90	2701,60
HTML DINAM.	(%)		
Frecuencia	54,44		
Bytes	2,75		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	11,68	22,28	3,84
Mediana	4,11	8,54	0,52
Desv. Típica	22,92	38,13	71,22
Coef. var. de Pearson	1,96	1,71	18,55
Mínimo	0,21	0,00	0,01
Máximo	996,96	583,91	4404,70
TEXTO	(%)		
Frecuencia	0,38		
Bytes	2,70		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	61,70	115,90	10,12
Mediana	6,91	11,80	0,18
Desv. Típica	470,21	423,82	65,13
Coef. var. de Pearson	7,62	3,66	6,44
Mínimo	0,26	0,02	0,00
Máximo	9431,00	7603,80	950,39

HTTP	(%)		
Frecuencia	0,00		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,45	1,43	5,31
Mediana	0,39	1,41	0,40
Desv. Típica	0,19	1,38	7,74
Coef. var. de Pearson	0,43	0,96	1,46
Mínimo	0,32	0,02	0,15
Máximo	0,89	3,34	20,85
OTROS	(%)		
Frecuencia	0,80		
Bytes	1,06		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	11,09	160,42	3,08
Mediana	1,09	28,19	0,06
Desv. Típica	191,25	430,21	67,49
Coef. var. de Pearson	17,24	2,68	21,88
Mínimo	0,29	0,02	0,00
Máximo	8841,30	4198,70	2313,00
SCRIPT	(%)		
Frecuencia	1,41		
Bytes	0,68		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,57	51,65	0,91
Mediana	3,26	14,52	0,23
Desv. Típica	9,97	157,80	8,48
Coef. var. de Pearson	1,52	3,06	9,33
Mínimo	0,31	0,01	0,00
Máximo	167,90	4546,50	283,66
DOCUMENTOS	(%)		
Frecuencia	0,06		
Bytes	0,28		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	44,62	28,13	3,29
Mediana	23,15	17,88	0,60
Desv. Típica	165,89	32,73	6,96
Coef. var. de Pearson	3,72	1,16	2,12
Mínimo	0,43	0,09	0,04
Máximo	1891,40	197,69	44,86

IRCACHE	(%)		
Frecuencia	0,21		
Bytes	27,05		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	581,31	88,23	106,81
Mediana	1025,90	71,73	8,17
Desv. Típica	458,05	96,76	372,38
Coef. var. de Pearson	0,79	1,10	3,49
Mínimo	36,10	0,06	0,29
Máximo	1025,90	918,42	4148,00
PRESENTACION	(%)		
Frecuencia	1,04		
Bytes	0,31		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	4,57	29,48	2,08
Mediana	2,75	10,14	0,28
Desv. Típica	5,36	78,51	20,33
Coef. var. de Pearson	1,17	2,66	9,78
Mínimo	0,31	0,00	0,00
Máximo	51,51	1987,80	610,23
DESCONOCIDOS	(%)		
Frecuencia	0,95		
Bytes	14,33		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	155,25	21,07	58,14
Mediana	1,96	6,92	0,26
Desv. Típica	673,28	38,36	222,59
Coef. var. de Pearson	4,34	1,82	3,83
Mínimo	0,17	0,01	0,01
Máximo	5124,40	513,85	1653,80

Tabla 3.25: Estudio por clase de objeto de las trazas del 23/03/04

IMÁGENES	(%)		
Frecuencia	54,42		
Bytes	34,72		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	13,66	39,87	1,73
Mediana	3,29	13,58	0,26
Desv. Típica	54,45	123,77	16,50
Coef. var. de Pearson	3,99	3,10	9,52
Mínimo	0,02	0,00	0,00
Máximo	9694,50	6662,10	2027,20
COMPRIMIDOS	(%)		
Frecuencia	0,16		
Bytes	1,76		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	126,74	82,54	61,55
Mediana	1,48	26,19	0,15
Desv. Típica	477,65	153,83	153,39
Coef. var. de Pearson	3,77	1,86	2,49
Mínimo	0,71	0,15	0,00
Máximo	4211,10	1397,50	1043,30
APLICACIÓN	(%)		
Frecuencia	0,75		
Bytes	5,55		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	73,60	41,08	44,85
Mediana	1,27	20,61	0,06
Desv. Típica	445,98	68,27	349,44
Coef. var. de Pearson	6,06	1,66	7,79
Mínimo	0,22	0,04	0,02
Máximo	7,65	651,20	10800,00
VIDEO	(%)		
Frecuencia	0,14		
Bytes	8,50		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	795,99	61,04	61,72
Mediana	640,35	49,99	8,61
Desv. Típica	964,10	45,42	259,30
Coef. var. de Pearson	1,21	0,74	4,20
Mínimo	0,34	0,00	0,08
Máximo	7866,60	163,15	2310,40

SONIDO	(%)		
Frecuencia	0,11		
Bytes	1,01		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	208,02	28,52	38,99
Mediana	27,82	21,84	0,64
Desv. Típica	745,96	29,88	132,45
Coef. var. de Pearson	3,59	1,05	3,40
Mínimo	1,45	0,07	0,07
Máximo	6266,20	193,40	1050,50
HTML	(%)		
Frecuencia	9,62		
Bytes	7,81		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	14,32	25,24	3,44
Mediana	5,35	11,62	0,57
Desv. Típica	26,67	108,14	37,82
Coef. var. de Pearson	1,86	4,28	10,98
Mínimo	0,21	0,00	0,00
Máximo	692,84	7705,60	1967,20
HTML DINAM.	(%)		
Frecuencia	27,43		
Bytes	1,91		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,05	16,68	2,24
Mediana	0,78	11,43	0,21
Desv. Típica	15,16	22,81	34,75
Coef. var. de Pearson	2,51	1,39	15,51
Mínimo	0,22	0,01	0,01
Máximo	468,88	396,83	1954,30
TEXTO	(%)		
Frecuencia	0,46		
Bytes	2,43		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	75,60	223,59	6,66
Mediana	2,18	10,97	0,17
Desv. Típica	376,03	697,03	72,61
Coef. var. de Pearson	4,97	3,12	11,36
Mínimo	0,23	0,02	0,00
Máximo	5768,90	5209,80	1748,00

HTTP	(%)		
Frecuencia	0,01		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,43	1,86	16,85
Mediana	0,33	1,72	0,17
Desv. Típica	0,26	2,22	58,05
Coef. var. de Pearson	0,60	1,20	3,45
Mínimo	0,28	0,00	0,15
Máximo	1,19	8,12	248,43
OTROS	(%)		
Frecuencia	0,83		
Bytes	1,68		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	30,15	48,00	2,92
Mediana	0,99	18,41	0,08
Desv. Típica	522,64	138,14	23,32
Coef. var. de Pearson	17,34	2,88	8,00
Mínimo	0,25	0,08	0,00
Máximo	16368,00	2099,40	347,09
SCRIPT	(%)		
Frecuencia	2,04		
Bytes	0,46		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,80	63,32	1,03
Mediana	3,12	16,09	0,21
Desv. Típica	10,37	217,33	9,93
Coef. var. de Pearson	1,53	3,43	9,68
Mínimo	0,35	0,00	0,00
Máximo	126,04	4088,20	340,17
DOCUMENTOS	(%)		
Frecuencia	0,15		
Bytes	0,44		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	53,76	28,98	12,41
Mediana	26,07	28,80	0,65
Desv. Típica	231,03	24,21	55,67
Coef. var. de Pearson	4,30	0,84	4,49
Mínimo	0,37	0,14	0,03
Máximo	2838,40	125,33	436,04

IRCACHE	(%)		
Frecuencia	0,41		
Bytes	31,52		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	589,29	94,67	109,66
Mediana	1025,90	76,19	7,83
Desv. Típica	458,67	100,53	353,00
Coef. var. de Pearson	0,78	1,06	353,76
Mínimo	40,10	0,05	3,23
Máximo	1025,90	884,38	4030,40
PRESENTACION	(%)		
Frecuencia	1,72		
Bytes	0,28		
	4,65	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	4,65	31,92	4,48
Mediana	2,61	10,07	0,29
Desv. Típica	5,69	102,00	75,87
Coef. var. de Pearson	1,22	3,20	16,92
Mínimo	0,22	0,00	0,00
Máximo	52,30	2412,90	1799,40
DESCONOCIDOS	(%)		
Frecuencia	1,74		
Bytes	1,94		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	15,48	28,54	11,24
Mediana	1,96	9,21	0,21
Desv. Típica	98,29	41,20	94,96
Coef. var. de Pearson	6,35	1,44	8,45
Mínimo	0,13	0,00	0,01
Máximo	1268,00	435,50	2126,10

Tabla 3.26: Estudio por clase de objeto de las trazas del 24/03/04

IMÁGENES	(%)		
Frecuencia	73,70		
Bytes	22,32		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,95	40,23	1,80
Mediana	3,37	14,69	0,25
Desv. Típica	22,75	134,49	29,50
Variabilidad	2,86	3,34	16,37
Mínimo	0,22	0,00	0,00
Máximo	1806,60	15157,00	2366,40
COMPRESION	(%)		
Frecuencia	0,82		
Bytes	2,86		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	38,89	163,08	15,07
Mediana	1,40	92,65	0,02
Desv. Típica	500,63	213,16	124,45
Variabilidad	12,87	1,31	8,26
Mínimo	0,45	0,06	0,00
Máximo	15813,00	2600,30	2017,40
APLICACIÓN	(%)		
Frecuencia	0,84		
Bytes	5,09		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	79,83	55,89	27,28
Mediana	2,36	23,16	0,16
Desv. Típica	545,07	83,79	153,82
Variabilidad	6,83	1,50	5,64
Mínimo	0,20	0,02	0,01
Máximo	10188,00	732,91	2488,10
VIDEO	(%)		
Frecuencia	0,25		
Bytes	5,87		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	317,12	30,31	199,49
Mediana	60,85	0,83	59,17
Desv. Típica	572,46	46,96	338,96
Variabilidad	1,81	1,55	1,70
Mínimo	0,26	0,05	0,09
Máximo	5765,10	158,12	3967,00

SONIDO	(%)		
Frecuencia	0,14		
Bytes	0,94		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	284,29	113,35	67,38
Mediana	33,71	12,64	1,41
Desv. Típica	981,83	418,40	232,74
Variabilidad	3,45	3,69	3,45
Mínimo	0,86	0,13	0,00
Máximo	6154,30	2476,20	1503,30
HTML	(%)		
Frecuencia	12,15		
Bytes	7,09		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	14,00	30,48	4,91
Mediana	7,08	13,01	0,48
Desv. Típica	20,43	109,32	44,55
Variabilidad	1,46	3,59	9,08
Mínimo	0,21	0,00	0,00
Máximo	631,51	6811,50	1995,30
HTML DINAM.	(%)		
Frecuencia	3,00		
Bytes	1,90		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	11,83	12,39	5,29
Mediana	2,67	4,01	0,53
Desv. Típica	19,57	20,61	46,22
Variabilidad	1,65	1,66	8,73
Mínimo	0,22	0,00	0,01
Máximo	236,88	252,72	1992,10
TEXTO	(%)		
Frecuencia	0,73		
Bytes	2,42		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	45,02	63,67	6,00
Mediana	0,43	2,53	0,17
Desv. Típica	209,77	382,00	37,33
Variabilidad	4,66	6,00	6,23
Mínimo	0,23	0,05	0,00
Máximo	3400,10	6066,20	904,44

HTTP	(%)		
Frecuencia	0,01		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,74	2,28	99,75
Mediana	0,39	0,10	3,37
Desv. Típica	1,02	3,85	197,06
Variabilidad	1,38	1,69	1,98
Mínimo	0,30	0,00	0,15
Máximo	4,18	0,00	671,69
OTROS	(%)		
Frecuencia	1,13		
Bytes	3,54		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	62,89	45,75	20,99
Mediana	1,00	14,53	0,09
Desv. Típica	1013,40	141,75	257,63
Variabilidad	16,11	3,10	12,27
Mínimo	0,29	0,07	0,00
Máximo	26616,00	1678,10	5565,10
SCRIPT	(%)		
Frecuencia	2,54		
Bytes	0,38		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,35	45,68	1,28
Mediana	2,56	13,55	0,22
Desv. Típica	10,04	124,27	8,17
Variabilidad	1,58	2,72	6,40
Mínimo	0,34	0,07	0,00
Máximo	129,88	2337,80	141,91
DOCUMENTOS	(%)		
Frecuencia	0,16		
Bytes	0,22		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	28,62	70,44	13,56
Mediana	21,91	17,83	1,05
Desv. Típica	41,04	517,04	65,47
Variabilidad	1,43	7,34	4,83
Mínimo	0,51	0,38	0,01
Máximo	327,87	6170,90	567,46

IRCACHE	(%)		
Frecuencia	0,53		
Bytes	31,52		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	591,58	90,67	96,43
Mediana	1025,90	76,19	7,83
Desv. Típica	457,62	88,42	374,59
Variabilidad	0,77	0,98	3,88
Mínimo	40,10	0,04	0,29
Máximo	1025,90	876,07	4420,60
PRESENTACION	(%)		
Frecuencia	2,22		
Bytes	0,25		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	4,34	24,54	1,35
Mediana	2,48	8,04	0,32
Desv. Típica	6,57	59,13	5,19
Variabilidad	1,51	2,41	10,69
Mínimo	0,22	0,03	0,00
Máximo	95,04	965,82	67,90
DESCONOCIDOS	(%)		
Frecuencia	1,77		
Bytes	15,61		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	174,33	15,05	60,19
Mediana	1,87	6,58	0,26
Desv. Típica	730,62	29,89	224,40
Variabilidad	4,19	1,99	3,73
Mínimo	0,20	0,02	0,00
Máximo	5279,20	664,77	2085,40

Tabla 3.27: Estudio por clase de objeto de las trazas del 25/03/04

IMÁGENES	(%)		
Frecuencia	72,39		
Bytes	18,71		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,36	41,88	1,38
Mediana	2,90	14,36	0,24
Desv. Típica	36,05	138,37	21,13
Variabilidad	4,90	3,30	15,30
Mínimo	0,22	0,00	0,00
Máximo	3392,00	9735,40	1926,70
COMPRESION	(%)		
Frecuencia	2,04		
Bytes	13,93		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	82,72	194,65	7,95
Mediana	1,40	127,04	0,01
Desv. Típica	1063,20	260,83	106,61
Variabilidad	12,85	1,34	13,42
Mínimo	0,45	0,01	0,01
Máximo	44771,00	3907,50	4516,90
APLICACIÓN	(%)		
Frecuencia	1,01		
Bytes	19,20		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	409,17	42,53	318,94
Mediana	2,47	13,98	0,18
Desv. Típica	2609,70	101,66	1331,30
Variabilidad	6,38	2,39	4,17
Mínimo	0,02	0,02	0,02
Máximo	45907,00	1187,50	10800,00
VIDEO	(%)		
Frecuencia	0,16		
Bytes	5,41		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	671,42	47,15	99,65
Mediana	2547734,00	19,31	7,73
Desv. Típica	2312,60	50,47	428,73
Variabilidad	3,44	1,07	4,30
Mínimo	0,47	0,09	0,06
Máximo	29211,00	184,43	3764,30

SONIDO	(%)		
Frecuencia	0,50		
Bytes	0,75		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	113,25	43,30	111,24
Mediana	25,83	7,47	1,36
Desv. Típica	284,97	90,90	273,35
Variabilidad	2,52	2,10	2,46
Mínimo	0,34	0,03	0,01
Máximo	2739,60	669,30	2277,90
HTML	(%)		
Frecuencia	11,10		
Bytes	5,72		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	14,00	24,09	3,85
Mediana	6,54	10,77	0,54
Desv. Típica	19,73	74,01	33,22
Variabilidad	1,41	3,07	8,62
Mínimo	0,22	0,00	0,00
Máximo	361,13	3297,90	1800,10
HTML DINAM.	(%)		
Frecuencia	3,20		
Bytes	2,24		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	13,17	17,38	5,20
Mediana	4,17	7,27	0,60
Desv. Típica	98,12	28,75	73,98
Variabilidad	7,45	1,65	14,23
Mínimo	0,24	0,01	0,01
Máximo	5683,80	385,96	3391,70
TEXTO	(%)		
Frecuencia	0,60		
Bytes	3,09		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	95,28	56,09	3,76
Mediana	1,31	9,67	0,20
Desv. Típica	849,77	154,36	19,09
Variabilidad	8,92	2,75	5,08
Mínimo	0,23	0,01	0,00
Máximo	19945,00	2802,20	333,59

HTTP	(%)		
Frecuencia	0,01		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,45	1,15	52,16
Mediana	0,34	0,30	3,03
Desv. Típica	0,25	1,41	120,62
Variabilidad	0,56	1,23	2,31
Mínimo	0,30	0,00	0,16
Máximo	1,19	3,62	426,33
OTROS	(%)		
Frecuencia	1,21		
Bytes	0,49		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	9,10	215,73	4,81
Mediana	8,24	38,31	0,07
Desv. Típica	41,97	443,58	63,63
Variabilidad	4,61	2,06	13,23
Mínimo	0,29	0,05	0,00
Máximo	1304,90	4197,30	1938,80
SCRIPT	(%)		
Frecuencia	2,57		
Bytes	0,40		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,34	40,68	1,18
Mediana	2,56	12,55	0,24
Desv. Típica	10,36	105,72	9,90
Variabilidad	1,63	2,60	8,41
Mínimo	0,32	0,02	0,00
Máximo	190,75	2273,30	256,00
DOCUMENTOS	(%)		
Frecuencia	0,14		
Bytes	0,26		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	48,96	28,16	20,64
Mediana	25,99	16,28	0,62
Desv. Típica	162,03	34,83	137,64
Variabilidad	3,31	1,24	6,67
Mínimo	0,38	0,02	0,02
Máximo	1509,00	246,24	1291,00

IRCACHE	(%)		
Frecuencia	0,50		
Bytes	26,47		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	570,10	88,92	110,77
Mediana	860,11	73,48	7,79
Desv. Típica	458,12	95,64	450,38
Variabilidad	0,80	1,08	4,07
Mínimo	0,56	0,03	0,25
Máximo	1025,90	962,36	5609,70
PRESENTACION	(%)		
Frecuencia	2,08		
Bytes	0,23		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	4,74	27,99	1,02
Mediana	2,75	8,85	0,31
Desv. Típica	5,79	70,60	3,32
Variabilidad	1,22	2,52	3,27
Mínimo	0,36	0,03	0,00
Máximo	78,24	1219,70	64,69
DESCONOCIDOS	(%)		
Frecuencia	2,49		
Bytes	3,09		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	20,70	22,47	5,22
Mediana	1,99	8,94	0,24
Desv. Típica	135,15	30,95	64,10
Variabilidad	6,53	1,38	12,28
Mínimo	0,20	0,01	0,01
Máximo	3577,00	307,35	2739,70

Tabla 3.28: Estudio por clase de objeto de las trazas del 26/03/04

IMÁGENES	(%)		
Frecuencia	71,41		
Bytes	18,23		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,62	40,76	1,10
Mediana	3,20	14,19	0,02
Desv. Típica	27,68	115,64	11,93
Variabilidad	3,63	2,84	10,88
Mínimo	0,26	0,01	0,00
Máximo	2544,90	7960,90	940,78
COMPRESION	(%)		
Frecuencia	1,02		
Bytes	6,32		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	80,27	198,83	17,55
Mediana	1,48	98,89	0,02
Desv. Típica	843,03	263,50	130,79
Variabilidad	10,50	1,33	7,45
Mínimo	0,82	0,01	0,00
Máximo	18441,00	1483,40	2043,20
APLICACIÓN	(%)		
Frecuencia	0,58		
Bytes	7,22		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	225,36	43,75	366,83
Mediana	2,34	26,09	0,13
Desv. Típica	1085,60	44,69	1522,50
Variabilidad	4,82	1,02	4,15
Mínimo	0,20	0,04	0,01
Máximo	18053,00	191,54	10800,00
VIDEO	(%)		
Frecuencia	0,24		
Bytes	15,22		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	1220,10	50,02	80,36
Mediana	256,35	34,13	7,68
Desv. Típica	2586,10	48,74	189,58
Variabilidad	2,12	0,97	2,36
Mínimo	0,39	0,12	0,03
Máximo	25713,00	197,95	1559,00

SONIDO	(%)		
Frecuencia	0,29		
Bytes	1,98		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	188,07	32,99	243,54
Mediana	21,48	0,37	119,18
Desv. Típica	822,81	81,98	384,86
Variabilidad	4,38	2,48	1,58
Mínimo	0,39	0,00	0,01
Máximo	6172,30	617,66	2306,50
HTML	(%)		
Frecuencia	13,31		
Bytes	6,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	14,66	27,05	3,52
Mediana	5,90	10,83	0,48
Desv. Típica	26,80	83,41	31,36
Variabilidad	1,83	3,08	8,90
Mínimo	0,22	0,00	0,00
Máximo	692,80	4480,60	1990,10
HTML DINAM.	(%)		
Frecuencia	4,19		
Bytes	2,08		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	7,56	20,01	1,51
Mediana	0,85	16,58	0,10
Desv. Típica	18,98	25,41	9,44
Variabilidad	2,51	1,27	6,26
Mínimo	0,21	0,01	0,02
Máximo	499,97	515,99	315,43
TEXTO	(%)		
Frecuencia	0,73		
Bytes	2,09		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	51,89	50,23	6,20
Mediana	1,39	9,71	0,19
Desv. Típica	194,29	217,97	35,84
Variabilidad	3,74	4,34	5,78
Mínimo	0,28	0,00	0,00
Máximo	2577,70	2796,50	477,69

HTTP	(%)		
Frecuencia	0,00		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,35	0,69	20,02
Mediana	0,33	0,08	3,94
Desv. Típica	0,04	1,26	34,72
Variabilidad	0,10	1,83	1,73
Mínimo	0,33	0,00	0,16
Máximo	0,40	2,57	72,04
OTROS	(%)		
Frecuencia	1,25		
Bytes	0,74		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	8,74	285,79	3,12
Mediana	8,38	44,56	0,05
Desv. Típica	32,47	606,51	64,28
Variabilidad	3,72	2,12	20,63
Mínimo	0,29	0,05	0,00
Máximo	526,20	4241,20	2352,70
SCRIPT	(%)		
Frecuencia	2,50		
Bytes	0,34		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	5,80	29,32	0,94
Mediana	2,55	10,89	0,26
Desv. Típica	8,80	53,61	3,74
Variabilidad	1,52	1,83	3,98
Mínimo	0,32	0,01	0,00
Máximo	96,85	521,00	65,90
DOCUMENTOS	(%)		
Frecuencia	0,13		
Bytes	0,21		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	40,55	25,67	9,02
Mediana	24,05	14,79	1,45
Desv. Típica	103,24	27,19	40,90
Variabilidad	2,55	1,06	4,54
Mínimo	0,71	0,38	0,03
Máximo	942,41	116,82	372,84

IRCACHE	(%)		
Frecuencia	0,66		
Bytes	34,63		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	596,32	87,08	85,74
Mediana	1025,90	72,10	8,11
Desv. Típica	457,68	92,89	316,63
Variabilidad	0,77	1,07	3,69
Mínimo	40,10	0,07	0,32
Máximo	1025,90	918,34	4643,80
PRESENTACION	(%)		
Frecuencia	1,87		
Bytes	0,22		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	4,72	27,19	1,41
Mediana	2,79	8,61	0,34
Desv. Típica	5,83	59,57	5,82
Variabilidad	1,23	2,19	4,12
Mínimo	0,22	0,02	0,00
Máximo	52,40	636,47	119,89
DESCONOCIDOS	(%)		
Frecuencia	1,82		
Bytes	4,72		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	51,22	25,51	8,00
Mediana	1,99	10,11	0,21
Desv. Típica	266,47	37,52	53,63
Variabilidad	5,20	1,47	6,70
Mínimo	0,20	0,01	0,02
Máximo	7442,90	789,82	1175,60

Tabla 3.29: Estudio por clase de objeto de las trazas del 27/03/04

IMÁGENES	(%)		
Frecuencia	74,78		
Bytes	19,32		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	6,96	43,01	0,82
Mediana	3,25	16,19	0,24
Desv. Típica	15,05	113,87	6,38
Variabilidad	2,16	2,65	7,81
Mínimo	0,22	0,00	0,00
Máximo	1128,30	7778,00	484,00
COMPRESION	(%)		
Frecuencia	0,89		
Bytes	3,36		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	63,57	172,20	12,85
Mediana	1,40	116,62	0,01
Desv. Típica	642,46	210,00	118,14
Variabilidad	10,11	1,22	9,19
Mínimo	0,45	0,13	0,00
Máximo	16157,00	1399,40	2129,60
APLICACIÓN	(%)		
Frecuencia	0,68		
Bytes	1,73		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	30,36	36,47	21,31
Mediana	1,59	24,04	0,05
Desv. Típica	212,50	44,99	156,79
Variabilidad	7,00	1,23	7,36
Mínimo	0,23	0,05	0,01
Máximo	4207,50	458,84	2001,90
VIDEO	(%)		
Frecuencia	0,23		
Bytes	12,86		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	795,10	75,91	23,57
Mediana	536,38	87,13	5,97
Desv. Típica	1553,20	48,87	126,54
Variabilidad	1,95	0,64	5,37
Mínimo	0,42	0,10	0,15
Máximo	19406,00	161,78	1967,60

SONIDO	(%)		
Frecuencia	0,17		
Bytes	1,83		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	367,26	111,13	84,52
Mediana	25,61	27,34	1,34
Desv. Típica	1252,90	397,53	247,85
Variabilidad	3,41	3,58	2,93
Mínimo	0,83	0,01	0,01
Máximo	10484,00	2622,50	1676,60
HTML	(%)		
Frecuencia	10,69		
Bytes	6,55		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	12,73	20,18	3,50
Mediana	4,50	8,52	0,45
Desv. Típica	22,76	45,39	35,60
Variabilidad	1,79	2,25	10,16
Mínimo	0,21	0,00	0,00
Máximo	721,66	1814,20	1908,30
HTML DINAM.	(%)		
Frecuencia	2,74		
Bytes	2,40		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	12,30	19,42	1,58
Mediana	3,38	7,59	0,51
Desv. Típica	21,18	31,12	8,10
Variabilidad	1,72	1,60	5,12
Mínimo	0,21	0,00	0,01
Máximo	530,72	383,47	241,50
TEXTO	(%)		
Frecuencia	0,55		
Bytes	7,17		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	265,12	38,91	22,95
Mediana	0,69	5,69	0,21
Desv. Típica	1183,90	196,27	171,24
Variabilidad	4,47	5,04	7,46
Mínimo	0,27	0,02	0,00
Máximo	13861,00	3859,40	2691,90

HTTP	(%)		
Frecuencia	0,00		
Bytes	0,00		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	0,33	0,12	2,83
Mediana	0,33	0,12	2,83
Desv. Típica	0,00	0,00	0,00
Variabilidad	0,00	0,00	0,00
Mínimo	0,33	0,12	2,83
Máximo	0,33	0,12	2,83
OTROS	(%)		
Frecuencia	0,86		
Bytes	0,87		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	17,39	265,49	1,11
Mediana	3,62	33,33	0,07
Desv. Típica	306,71	677,94	13,53
Variabilidad	17,63	2,55	12,14
Mínimo	0,29	0,04	0,00
Máximo	8575,40	4198,70	370,84
SCRIPT	(%)		
Frecuencia	2,38		
Bytes	0,37		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	5,71	38,65	0,87
Mediana	2,59	12,97	0,25
Desv. Típica	8,49	85,79	5,99
Variabilidad	1,49	2,22	6,91
Mínimo	0,35	0,01	0,00
Máximo	126,04	1128,10	134,35
DOCUMENTOS	(%)		
Frecuencia	0,16		
Bytes	0,26		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	25,22	24,77	2,42
Mediana	24,14	16,99	0,67
Desv. Típica	32,71	22,36	4,52
Variabilidad	1,30	0,90	1,87
Mínimo	0,54	1,20	0,05
Máximo	283,38	120,12	38,87

IRCACHE	(%)		
Frecuencia	0,60		
Bytes	36,31		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	566,21	84,00	111,22
Mediana	512,11	68,40	8,45
Desv. Típica	457,31	88,21	403,27
Variabilidad	0,81	1,05	3,63
Mínimo	2,27	0,05	0,31
Máximo	1025,90	893,62	4917,80
PRESENTACION	(%)		
Frecuencia	3,33		
Bytes	0,24		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	4,34	29,43	1,60
Mediana	2,62	7,80	0,35
Desv. Típica	5,23	134,90	7,69
Variabilidad	1,20	4,58	4,82
Mínimo	0,32	0,02	0,00
Máximo	51,41	3264,50	166,75
DESCONOCIDOS	(%)		
Frecuencia	1,95		
Bytes	6,72		
	TAMAÑOS EN KB	TASAS TRANS. (KB/sg)	TIEMPO RESP. (sg)
Media	49,17	25,65	4,69
Mediana	1,99	11,23	0,34
Desv. Típica	204,69	31,45	38,59
Variabilidad	4,16	1,23	8,23
Mínimo	0,20	0,01	0,01
Máximo	5070,50	484,54	970,44

Tabla 3.30: Estudio por clase de objeto de las trazas del 28/03/04

3.2.1.3. Resumen

Salvo pequeñas diferencias, ambos servidores proxy muestran el tráfico distribuido de la misma manera entre las distintas clases de objeto, siendo las imagenes, irCache, y las páginas estáticas las más importantes, formando más del 60% del tráfico.

En la figura 3.5 puede observarse además como las paginas estáticas aún siguen siendo mas populares que el uso de la tecnología dinamica.

La clase ‘desconocidos’ se encuentra en ambos servidores proxy formando más del 7% del tamaño de todas las peticiones. Esta clase esta formada por aquellos objetos cuyo tipo no estaba presente ni en el campo URL ni en el campo *content-type* de las trazas.

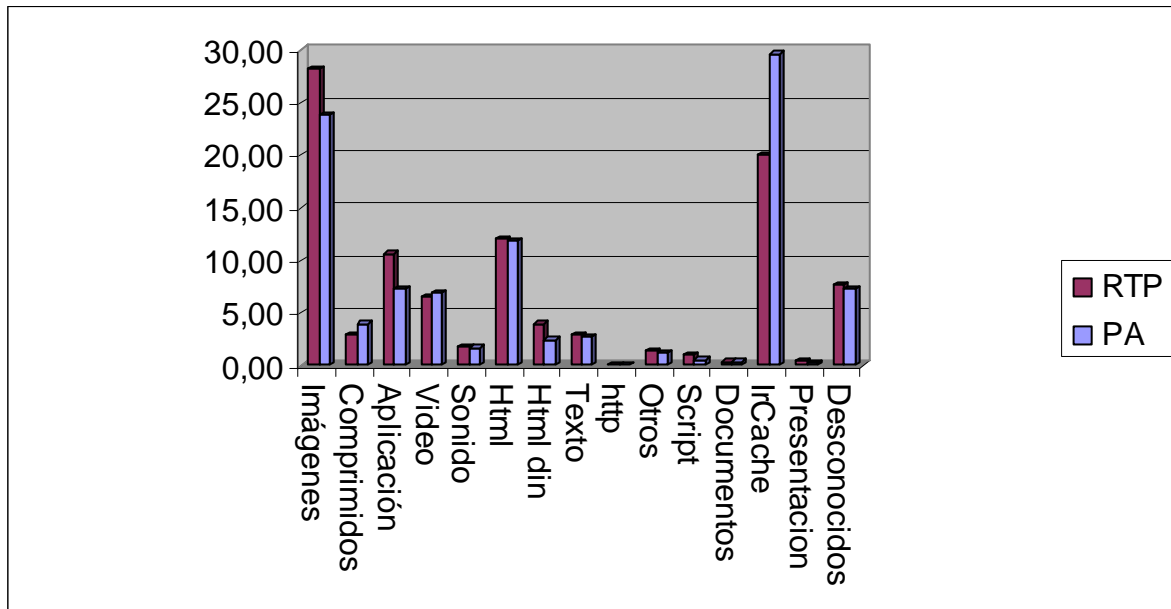


Figura 3.5. Porcentaje de bytes de las clases de objetos

El histograma de la figura 3.6 muestra que el tamaño de los objetos en PA son mucho mayores que en el caso RTP, con la única excepción de los objetos irCache. Estos objetos no pueden ser cacheados y sirven para el intercambio de información entre servidores proxy.

Los objetos de mayor tamaño son las imágenes y los videos, mientras que las clases que más aportan al tráfico (páginas estáticas e imágenes) no sobrepasan en ningun caso los 17 KB.

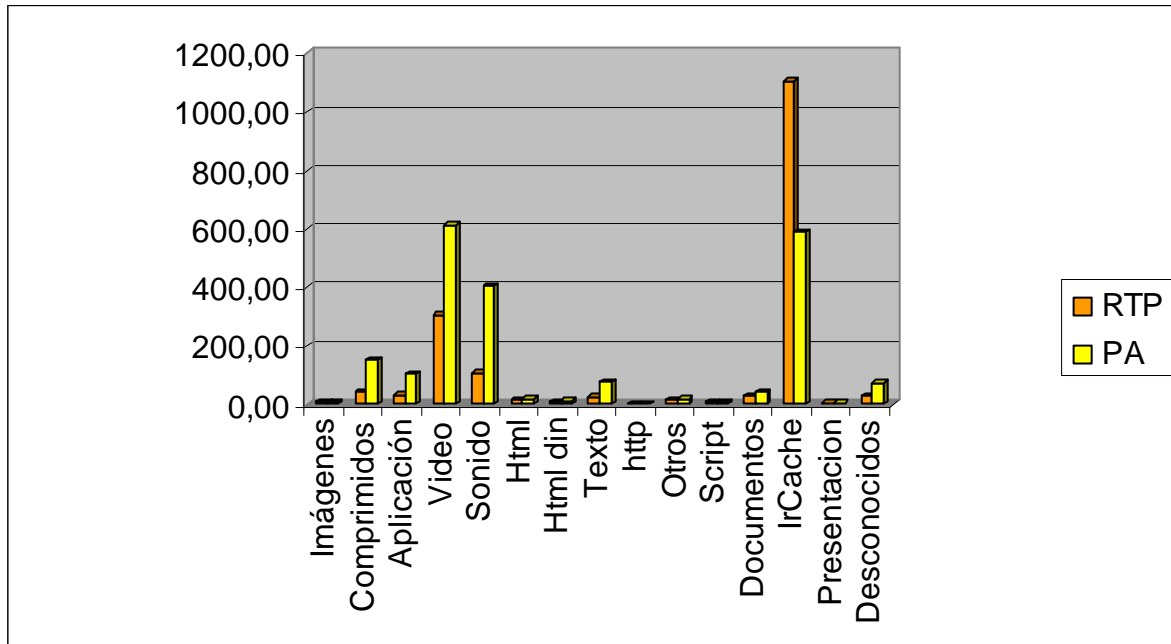


Figura 3.6. Tamaño medio de las clases de objetos

Como consecuencia de lo anterior, los tiempos de descarga son tambien mucho mayores que para el caso RTP, tal y como muestra el histograma de la figura 3.7

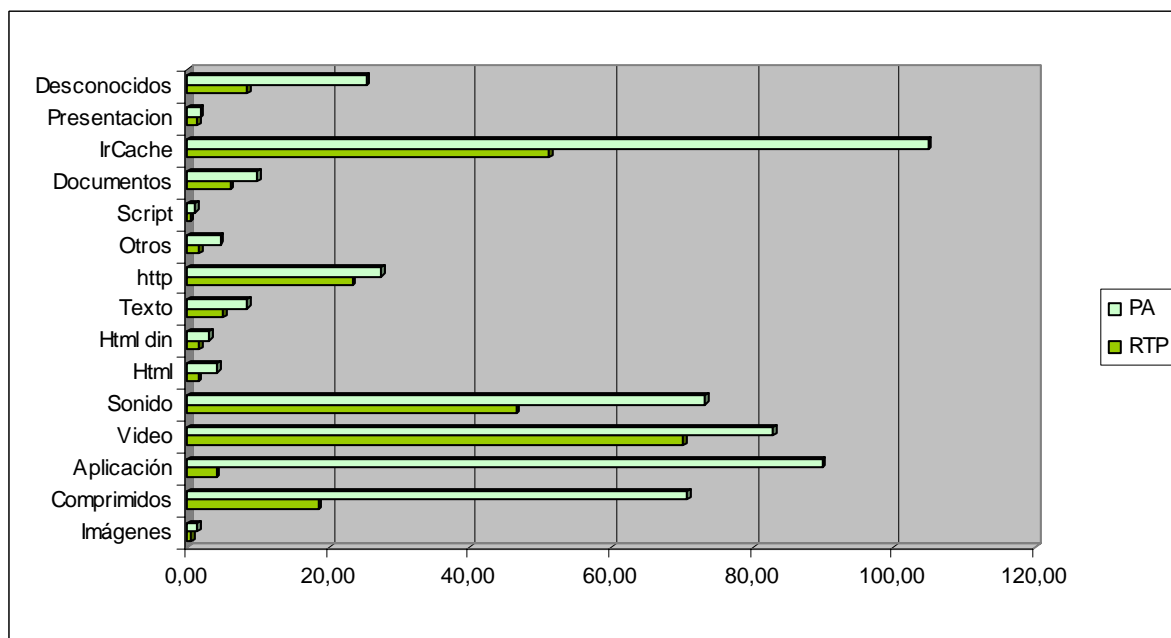


Figura 3.7. Tiempos de descarga de las clases de objeto

Por último, el histograma de la figura 3.8 muestra unas tasas de transferencia (medidas en KB/sg) mucho mayores para el servidor proxy RTP. Estas diferencias presentes para todas las clases de objetos, se acentúan más en el caso de los sonidos, las aplicaciones, y otros.

Las clases http y páginas dinamicas tienen las tasas de transferencia más bajas (no llegando a 8 KB/sg para el caso http ni a 28 KB/sg para el caso de las páginas dinamicas). Asi mismo comprimidos, sonidos, y otros tienen las más altas (llegando a más de 1135 KB/sg para el caso de otros en RTP).

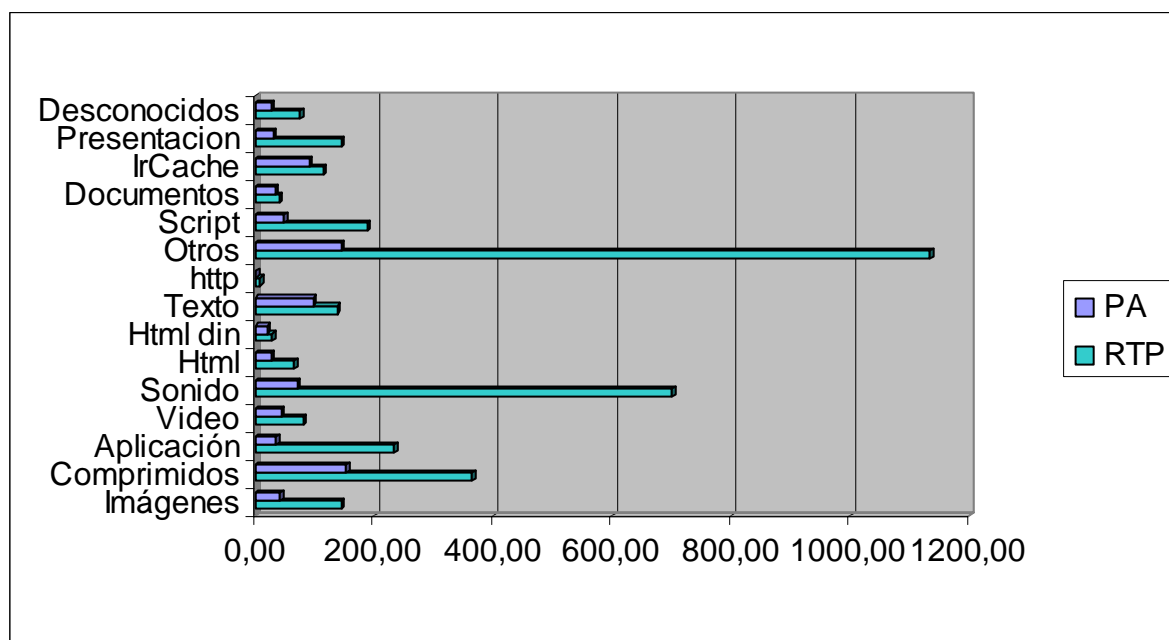


Figura 3.8. Histograma de comparación de las tasas de transferencia

3.2.2. ESTUDIO POR TIPO DE OBJETO

El estudio por tipo de objeto particulariza el estudio anterior centrándole únicamente en los tipos de objetos que más aportan al tráfico, entre todo ellos forman más del ochenta por ciento de las todas las peticiones.

Elegir un tipo de objeto representante para cada una de las clases de objeto empleadas en el apartado anterior, significaría reducir el volumen de datos, y de esta manera disminuir el tiempo y complejidad requeridos.

Estos son los tipos de objetos que estudiaremos a continuación:

- Jpg/Jpeg
- Gif
- Exe
- Html/Htm
- Zip
- Php
- Asp
- Avi
- Wav
- Irc

3.2.2.1. Servidor RTP

A continuación se detallan los resultados obtenidos para las trazas del servidor proxy RTP. Para cada uno de los tipos especificados anteriormente se muestran, el porcentaje de frecuencia con el que son referenciados, el porcentaje del tamaño que aportan al tráfico, y el tamaño medio en KB. Las tablas se encuentran ordenadas por el día al que pertenecen las trazas.

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES(%)	TAMAÑO MEDIO (KB)
Gif	46,64	8,33	3,61
Jpg/Jpeg	18,71	15,15	9,57
Exe	2,12	9,61	33,50
Html/Htm	10,65	12,14	15,13
Zip	0,69	2,27	29,46
Php	0,84	0,94	9,20
Asp	2,32	1,50	15,31
Avi	0,00	0,31	1120,70
Wav	0,01	0,04	125,54
Irc	0,44	18,31	1051,70

Tabla 3.31: Estudio de los principales tipos de objeto para las trazas del 22/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUECNA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	47,08	8,41	3,56
Jpg/Jpeg	19,01	16,18	10,09
Exe	1,50	9,17	46,57
Html/Htm	8,29	11,79	15,41
Zip	0,92	2,42	21,72
Php	0,79	0,79	8,76
Asp	0,96	1,40	15,14
Avi	0,00	0,15	392,83
Wav	0,02	0,04	71,18
Irc	0,87	17,65	1044,50

Tabla 3.32: Estudio de los principales tipos de objeto para las trazas del 23/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	46,28	8,98	3,58
Jpg/Jpeg	19,64	17,65	9,78
Exe	1,28	6,73	36,90
Html/Htm	10,89	12,38	15,08
Zip	0,58	3,85	72,60
Php	0,83	0,75	7,15
Asp	0,90	1,31	13,07
Avi	0,04	0,13	722,56
Wav	0,02	0,00	69,89
Irc	0,82	17,82	1061,90

Tabla 3.33: Estudio de los principales tipos de objeto para las trazas del 24/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	44,74	8,39	3,71
Jpg/Jpeg	19,08	17,62	10,44
Exe	2,71	7,12	18,89
Html/Htm	8,52	12,32	15,79
Zip	0,65	3,00	47,37
Php	0,87	0,87	8,48
Asp	0,92	1,70	18,38
Avi	0,00	1,79	18039,00
Wav	0,03	0,15	234,23
Irc	1,40	17,97	1097,90

Tabla 3.34: Estudio de los principales tipos de objeto para las trazas del 25/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	43,86	7,56	3,69
Jpg/Jpeg	19,11	14,62	9,61
Exe	1,75	13,18	55,67
Html/Htm	11,37	12,22	15,65
Zip	0,62	2,50	39,85
Php	0,86	0,84	7,87
Asp	1,12	1,66	14,88
Avi	0,00	0,16	1320,00
Wav	0,01	0,10	235,39
Irc	0,85	17,31	1108,90

Tabla 3.35: Estudio de los principales tipos de objeto para las trazas del 26/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	37,42	7,16	3,71
Jpg/Jpeg	19,69	15,24	9,50
Exe	1,85	5,13	20,46
Html/Htm	12,39	11,03	13,66
Zip	0,91	2,14	20,84
Php	1,34	1,12	6,92
Asp	1,27	1,90	13,41
Avi	0,06	0,61	72,22
Wav	0,01	0,25	182,84
Irc	1,70	28,74	1240,00

Tabla 3.36: Estudio de los principales tipos de objeto para las trazas del 27/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	34,82	6,48	3,66
Jpg/Jpeg	20,13	14,74	9,89
Exe	1,34	7,91	41,75
Html/Htm	8,77	9,12	12,45
Zip	0,52	0,58	16,99
Php	1,28	1,18	7,30
Asp	1,53	2,64	14,49
Avi	0,32	3,32	71,96
Wav	0,00	0,06	102,79
Irc	4,22	29,45	1232,50

Tabla 3.37: Estudio de los principales tipos de objeto para las trazas del 28/03/04

3.2.2.2. Servidor PA

Al igual que en el apartado anterior se detallan ahora los resultados obtenidos procedentes de las trazas del servidor proxy PA.

POR TIPO DE ARCHIVO			
TIPO	FRECUECNA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	19,89	5,79	3,66
Jpg/Jpeg	11,66	19,90	16,12
Exe	0,14	7,81	300,24
Html/Htm	7,32	20,50	22,17
Zip	0,06	2,48	464,03
Php	52,85	0,62	9,97
Asp	0,64	1,18	11,85
Avi	0,00	0,03	146,19
Wav	0,00	0,02	110,73
Irc	0,21	28,15	606,98

Tabla 3.38: Estudio de los principales tipos de objeto para las trazas del 22/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	20,32	5,77	3,80
Jpg/Jpeg	11,10	12,59	11,65
Exe	0,15	4,01	220,75
Html/Htm	6,99	13,05	18,85
Zip	0,25	0,99	32,63
Php	53,30	0,55	8,37
Asp	0,50	1,24	17,16
Avi	0,00	0,01	222,38
Wav	0,00	0,00	13,99
Irc	0,21	27,05	581,31

Tabla 3.39: Estudio de los principales tipos de objeto para las trazas del 23/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	32,62	4,86	3,87
Jpg/Jpeg	20,40	28,16	22,92
Exe	0,17	4,43	241,65
Html/Htm	9,46	7,72	14,56
Zip	0,13	1,44	134,43
Php	24,99	0,53	8,18
Asp	1,82	0,72	3,49
Avi	0,01	0,36	1132,00
Wav	0,01	0,06	133,35
Irc	0,41	31,52	589,29

Tabla 3.40: Estudio de los principales tipos de objeto para las trazas del 24/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	41,96	5,65	4,25
Jpg/Jpeg	29,79	15,14	10,69
Exe	0,48	4,39	97,51
Html/Htm	12,00	7,01	14,11
Zip	0,78	1,91	26,90
Php	1,08	0,74	12,99
Asp	1,13	0,55	8,47
Avi	0,00	0,02	228,04
Wav	0,01	0,00	26,13
Irc	0,53	31,52	591,58

Tabla 3.41: Estudio de los principales tipos de objeto para las trazas del 25/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	42,87	5,37	4,02
Jpg/Jpeg	27,67	11,63	10,06
Exe	0,65	16,21	476,34
Html/Htm	10,98	5,69	14,18
Zip	1,96	4,55	27,57
Php	1,08	0,85	12,29
Asp	1,12	0,56	13,70
Avi	0,00	0,08	554,60
Wav	0,01	0,03	74,89
Irc	0,50	26,47	570,10

Tabla 3.42: Estudio de los principales tipos de objeto para las trazas del 26/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	39,63	5,02	4,57
Jpg/Jpeg	29,78	11,51	9,25
Exe	0,30	6,77	293,32
Html/Htm	13,05	5,86	14,93
Zip	0,93	5,04	65,42
Php	0,98	0,59	9,29
Asp	2,04	0,49	3,20
Avi	0,01	0,03	261,13
Wav	0,01	0,00	21,46
Irc	0,66	34,63	596,32

Tabla 3.43: Estudio de los principales tipos de objeto para las trazas del 27/03/04

POR TIPO DE ARCHIVO			
TIPO	FRECUENCIA (%)	BYTES (%)	TAMAÑO MEDIO (KB)
Gif	44,23	5,43	4,25
Jpg/Jpeg	28,96	12,61	8,67
Exe	0,20	1,32	67,31
Html/Htm	10,55	6,48	12,89
Zip	0,86	1,77	34,61
Php	0,96	0,70	8,69
Asp	1,01	1,03	15,18
Avi	0,02	2,55	4439,80
Wav	0,01	0,01	191,82
Irc	0,60	36,31	566,21

Tabla 3.44: Estudio de los principales tipos de objeto para las trazas del 28/03/04

3.2.2.3. Resumen

Existe un gran número de peticiones para los objetos de tipo 'php' en el servidor PA para los tres primeros días del estudio (llegando a estar presentes en más del 50% de las peticiones, cuando la media no sobre pasa el 1%). Esta anomalía provoca una gran diferencia entre ambos servidores proxy para dicho objeto. Otras diferencias, mucho menores, también se encuentran en el caso de los 'exe' y los 'gif'.

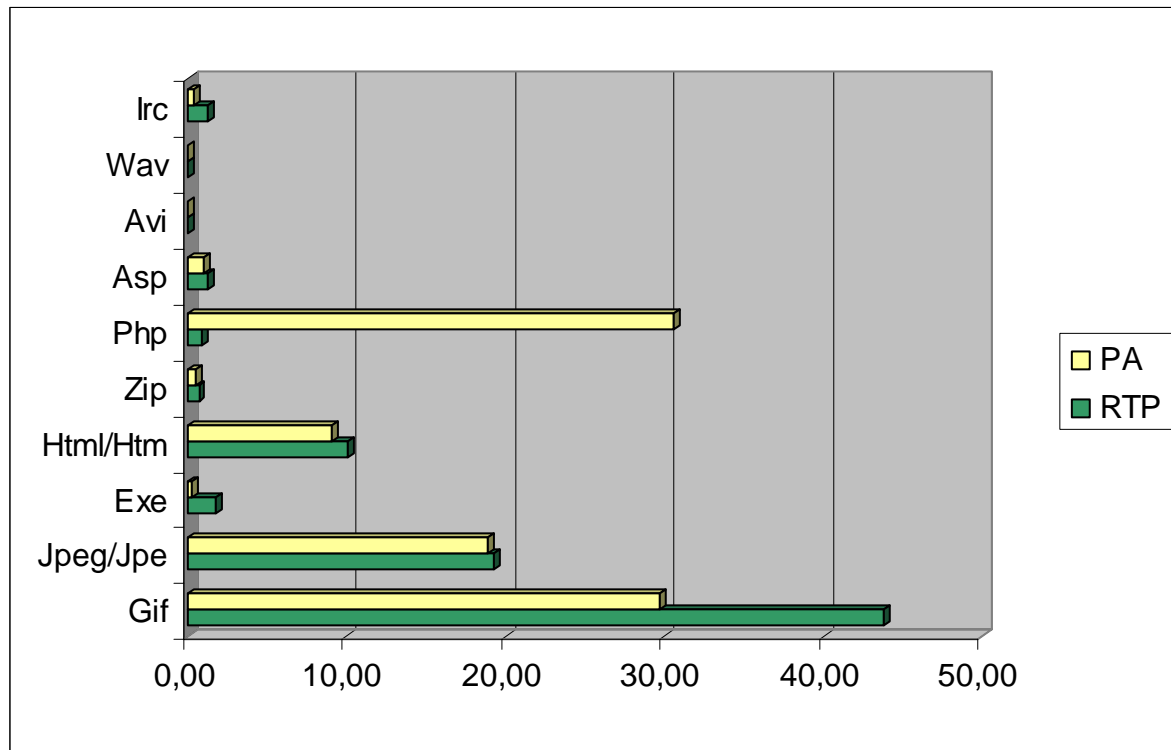


Figura 3.9. Frecuencias de peticiones de los principales tipos de objetos

Sin embargo, ese anormal número de peticiones explicada anteriormente para los objetos ‘php’, no se ve reflejado en el histograma de la figura 3.10, donde se muestra el porcentaje de bytes que aporta cada tipo al tráfico. En dicha gráfica tampoco se observan diferencias notables para los demás tipos a excepción del irCache que ya se explicó en el [apartado 3.2.1.3](#).

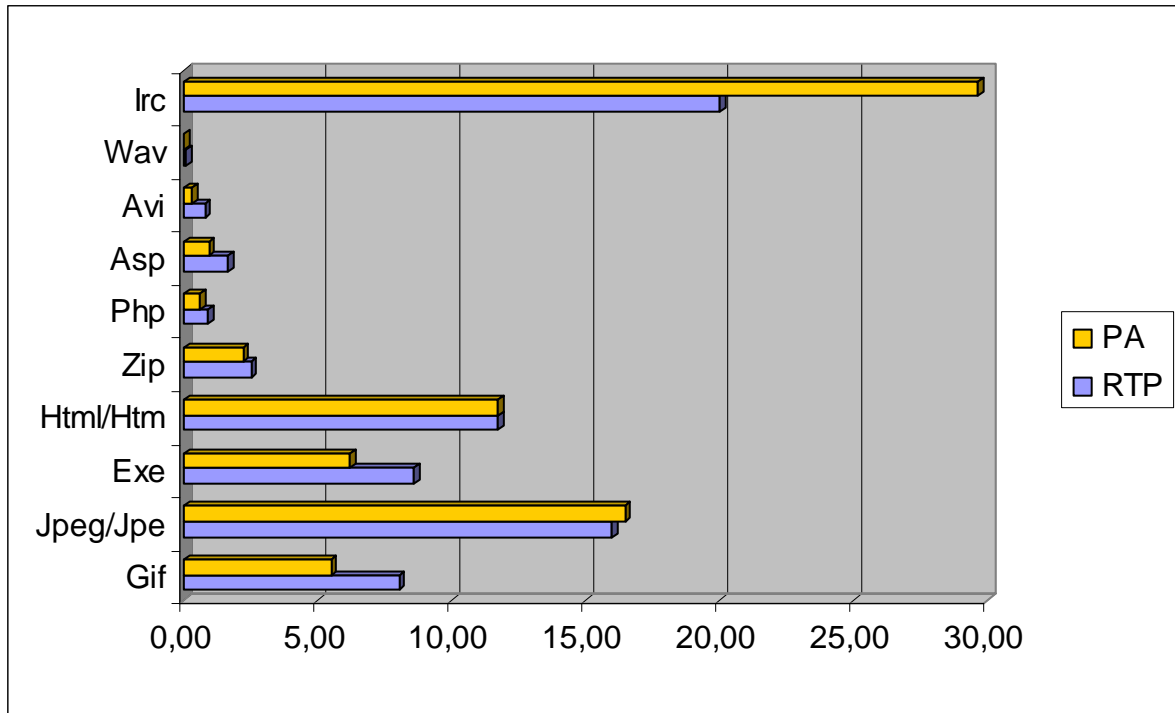


Figura 3.10. Porcentaje de bytes de los principales tipos de objeto

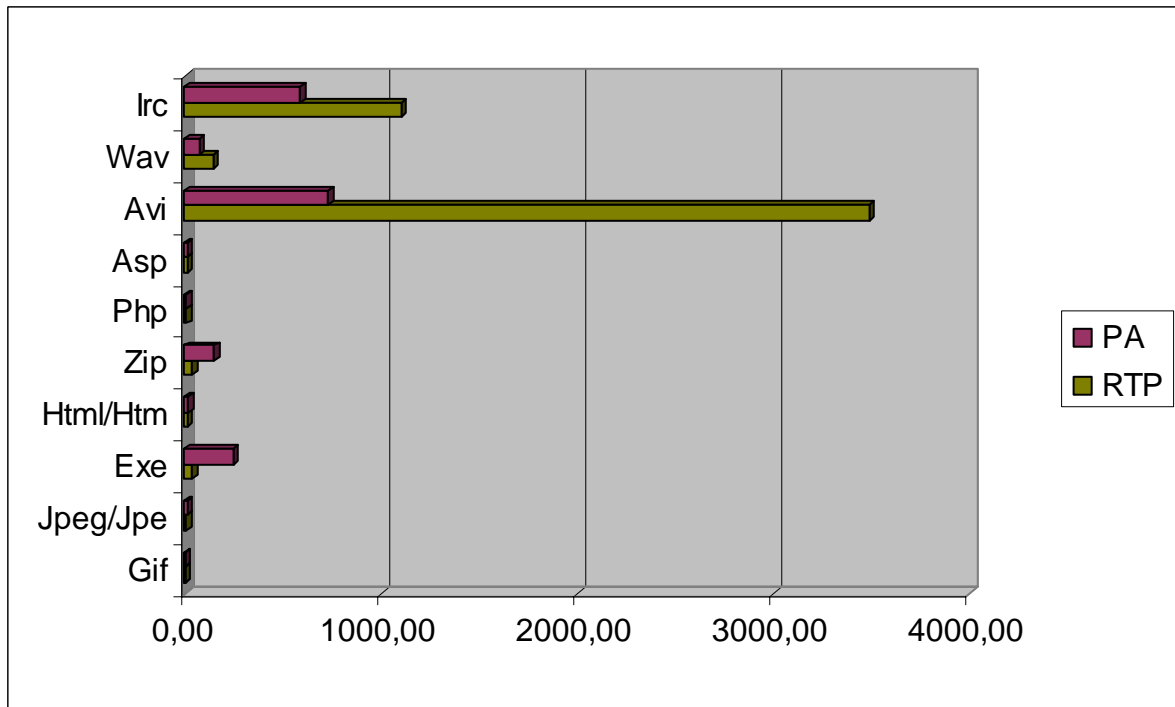


Figura 3.11. Tamaño medio de los principales tipos de objeto

CAPITULO 4: Caracterización del tráfico

En este capítulo se presentará un detallado análisis del comportamiento de peticiones a objetos, así como al tamaño de los mismos.

4.1. ESTUDIO DE SUCESIVAS PETICIONES

Estudiar la frecuencia con la que se solicitan los objetos más importantes podría permitir mejorar los algoritmos de reemplazo, determinar algunos parámetros de los proxy-cache, y en consecuencia mejorar el rendimiento de los mismos. Haciendo uso de la aplicación ‘distUrls2’, explicada en el [apartado 2.3](#), se estudia a continuación la distancia, tiempo, y tamaño que existen entre sucesivas peticiones a los objetos. Puesto que la aplicación utilizada para este estudio tiene un alto nivel de complejidad y consume gran cantidad de recursos del sistema, el estudio se centrará solo en los cuatro objetos que más peticiones recibieron en el servidor PA. Estos son:

1. *www.opasoft.com/work/scheduler.php*
2. *www.n3t.com.br/work/sscheduler.php*
3. *rad.msn.com/ADSAdClient31.dll*
4. *404.unterkunft-hotels-prag-tschechien.com/unterkunft-hotels-prag-tschechien.com.php*

4.1.1. DISTANCIAS ENTRE PETICIONES

Las gráficas que se exponen a continuación muestran las distancias existentes entre sucesivas peticiones a cada uno de los cuatro objetos anteriormente citados, esto es, representan el número de objetos solicitados entre sucesivas peticiones a un objeto dado respecto al tiempo.

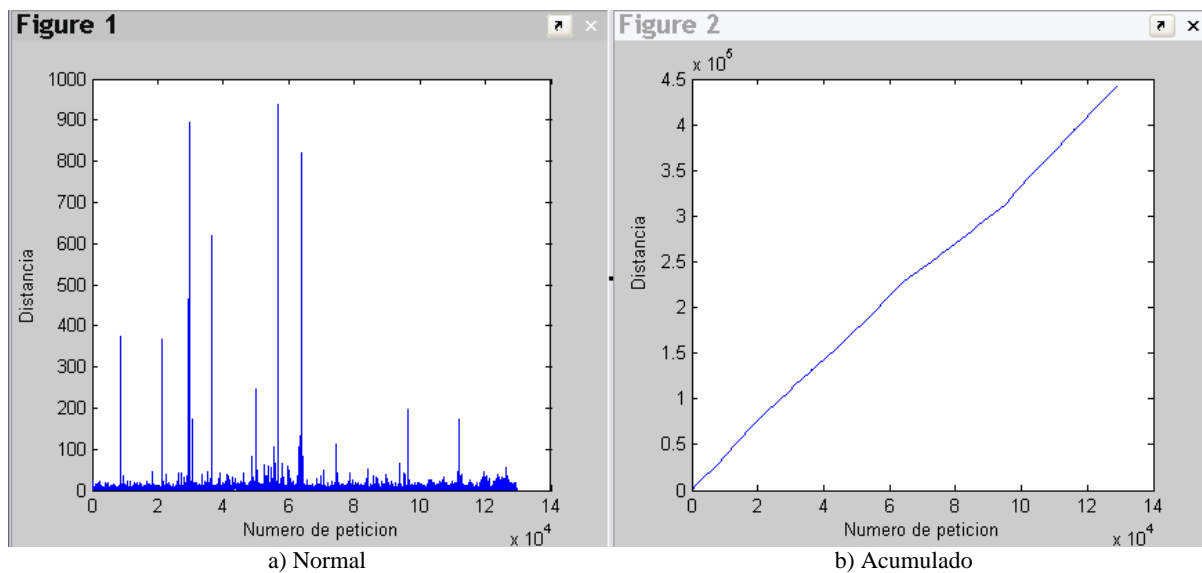


Figura 4.1. Distancias entre peticiones a 1 para el día 22/03/04

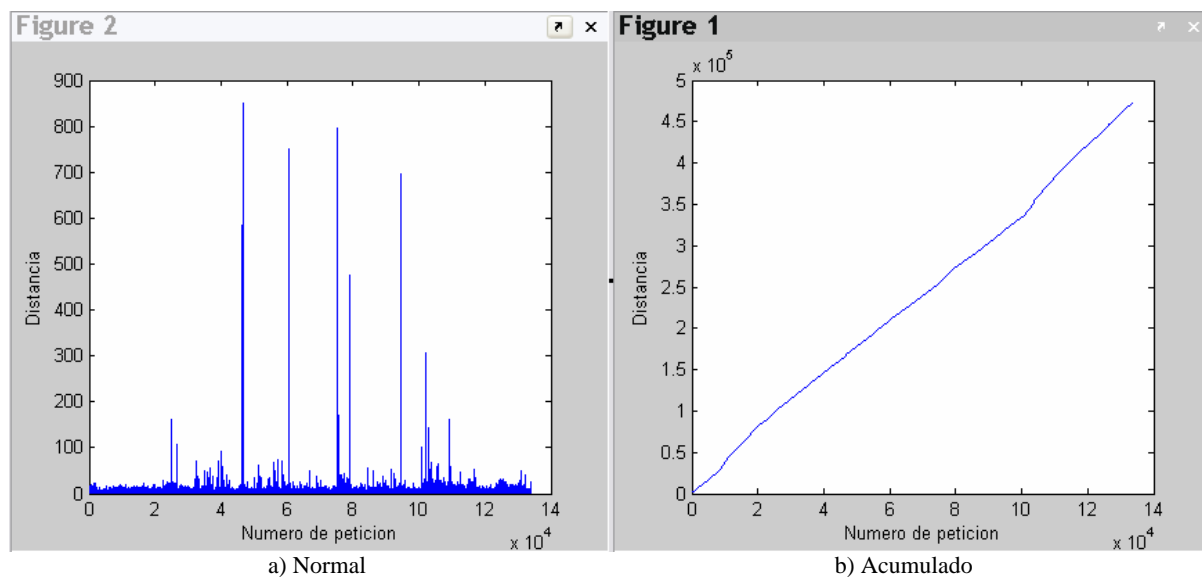


Figura 4.2. Distancias entre peticiones a 1 para el día 23/03/04

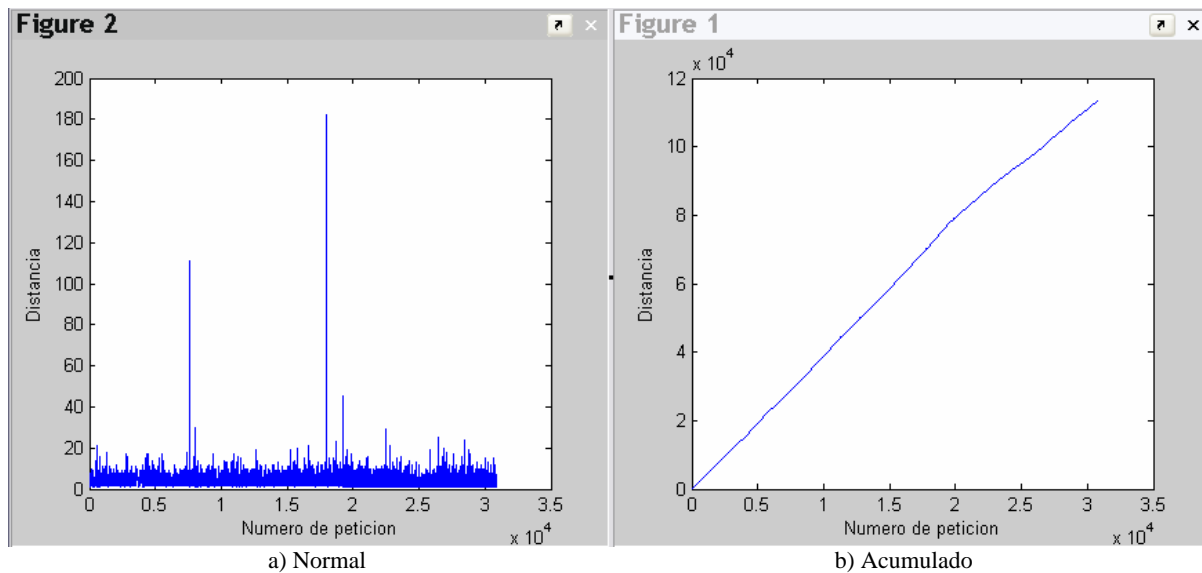


Figura 4.3. Distancias entre peticiones a 1 para el día 24/03/04

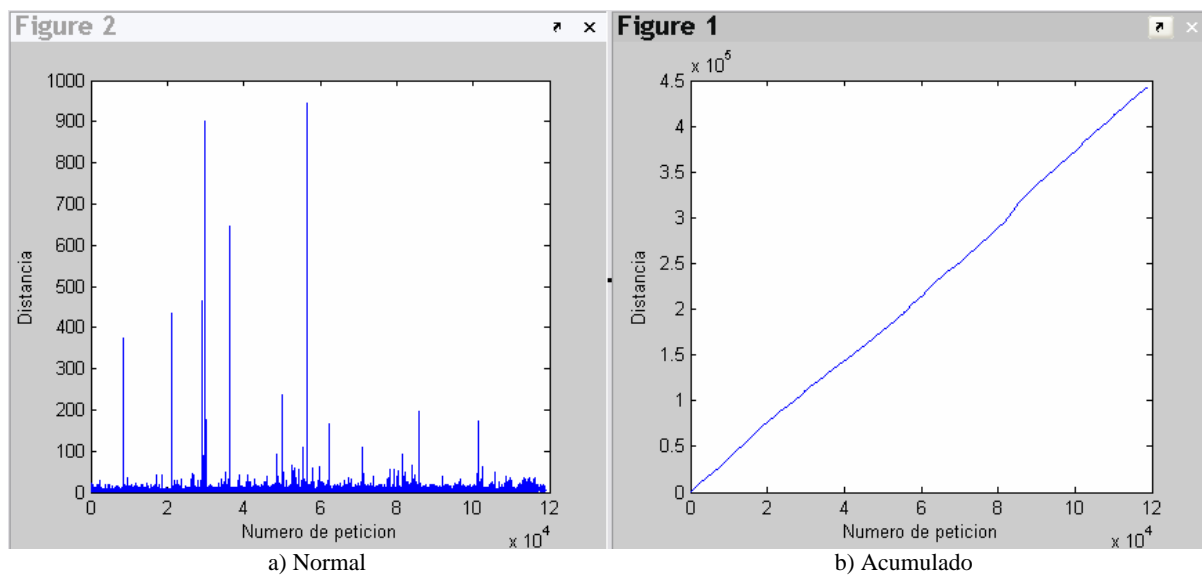


Figura 4.4. Distancias entre peticiones a 2 para el día 22/03/04

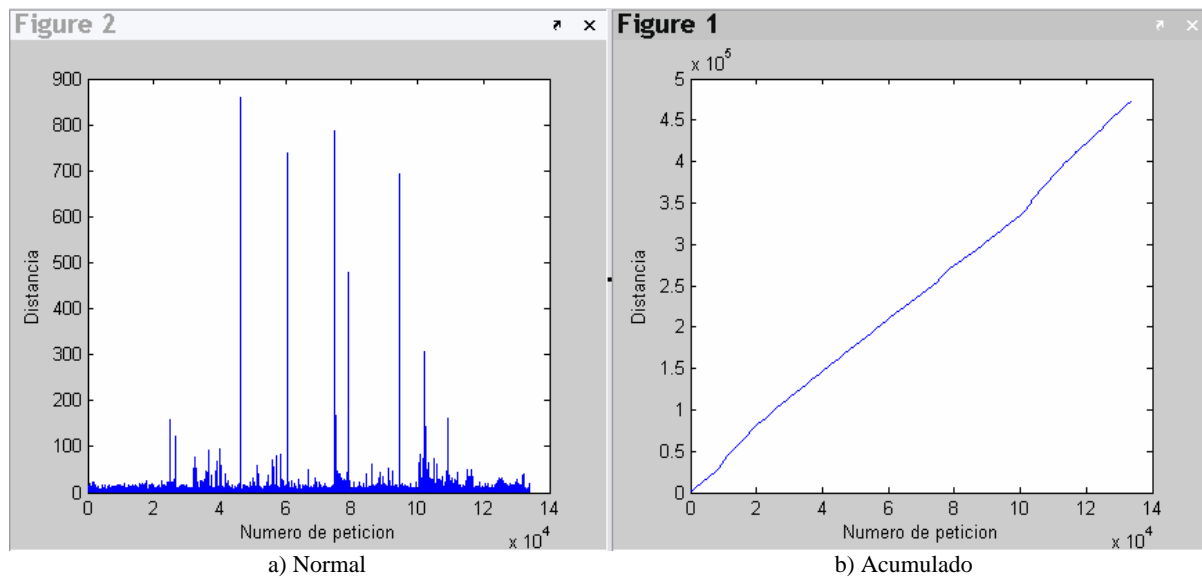


Figura 4.5. Distancias entre peticiones a 2 para el día 23/03/04

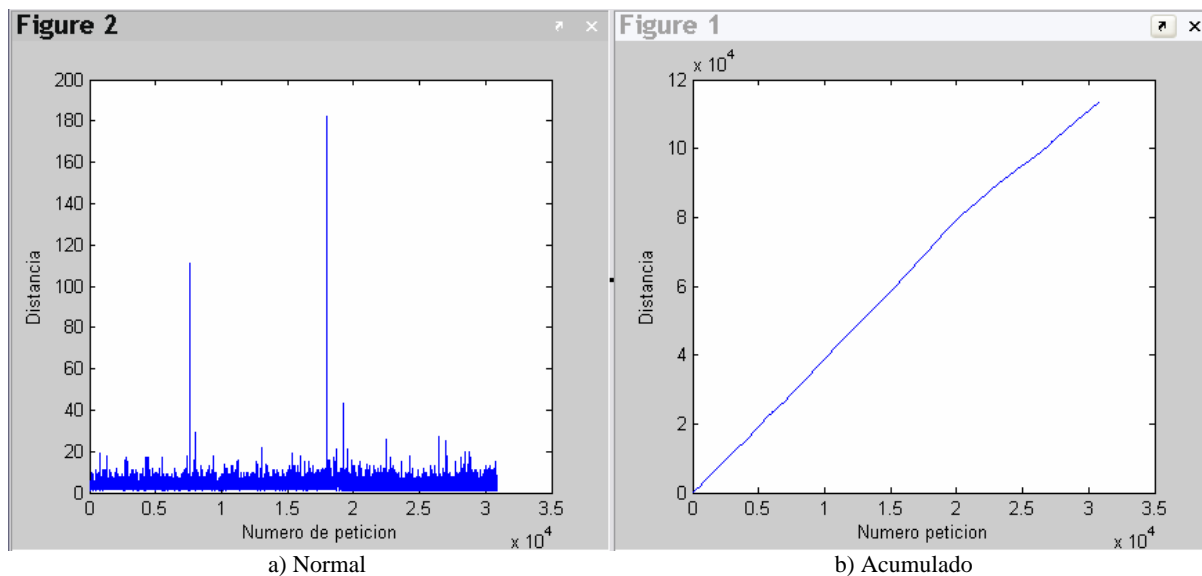


Figura 4.6. Distancias entre peticiones a 2 para el día 24/03/04

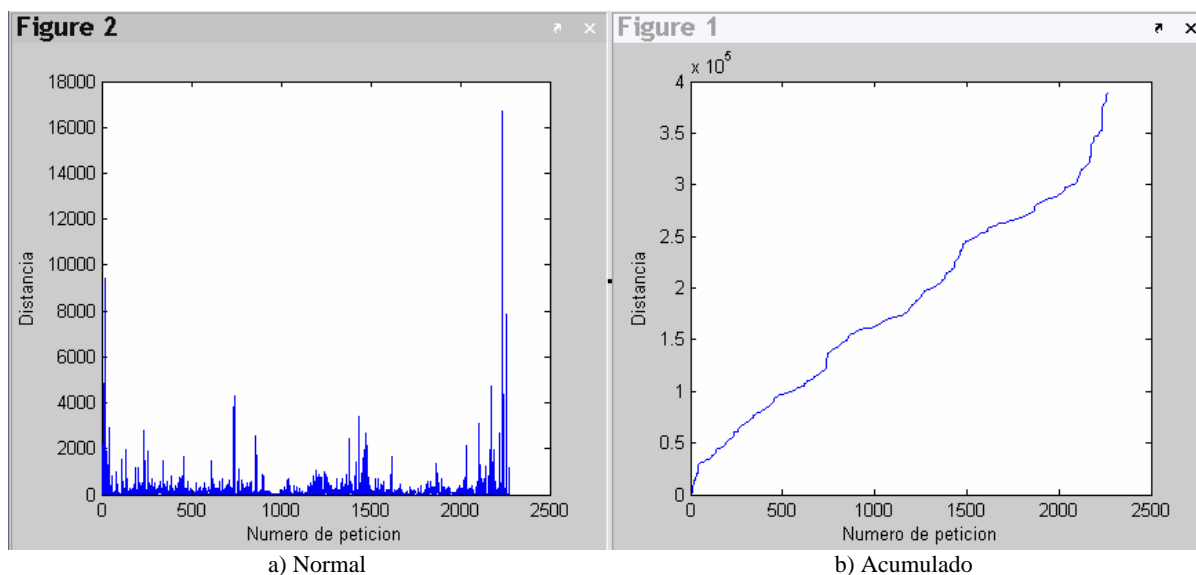


Figura 4.7. Distancias entre peticiones a 3 para el día 22/03/04

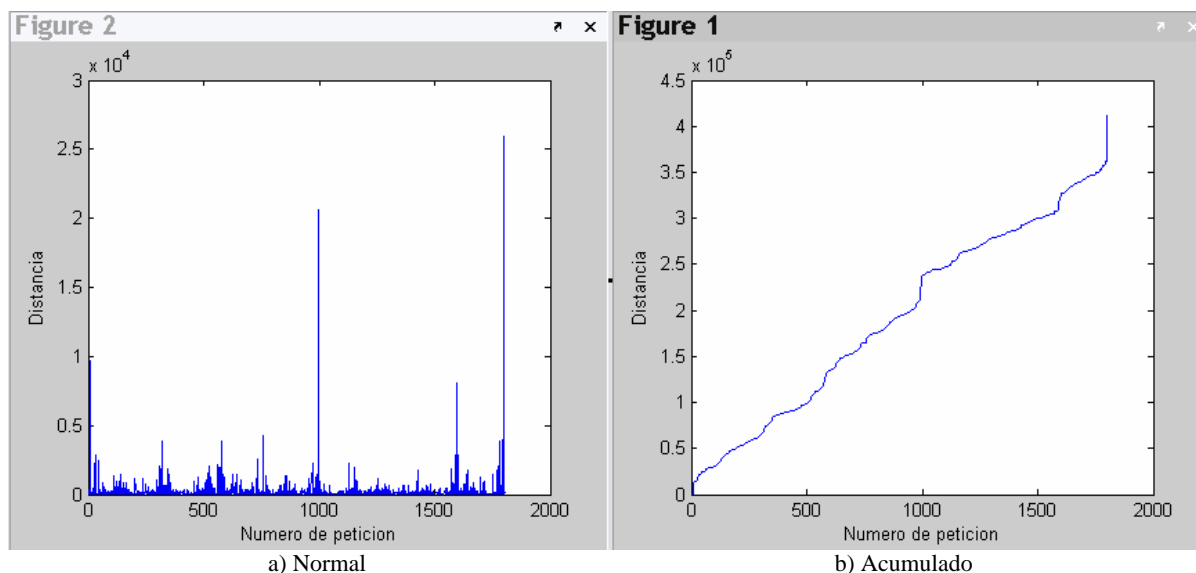


Figura 4.8. Distancias entre peticiones a 3 para el día 23/03/04

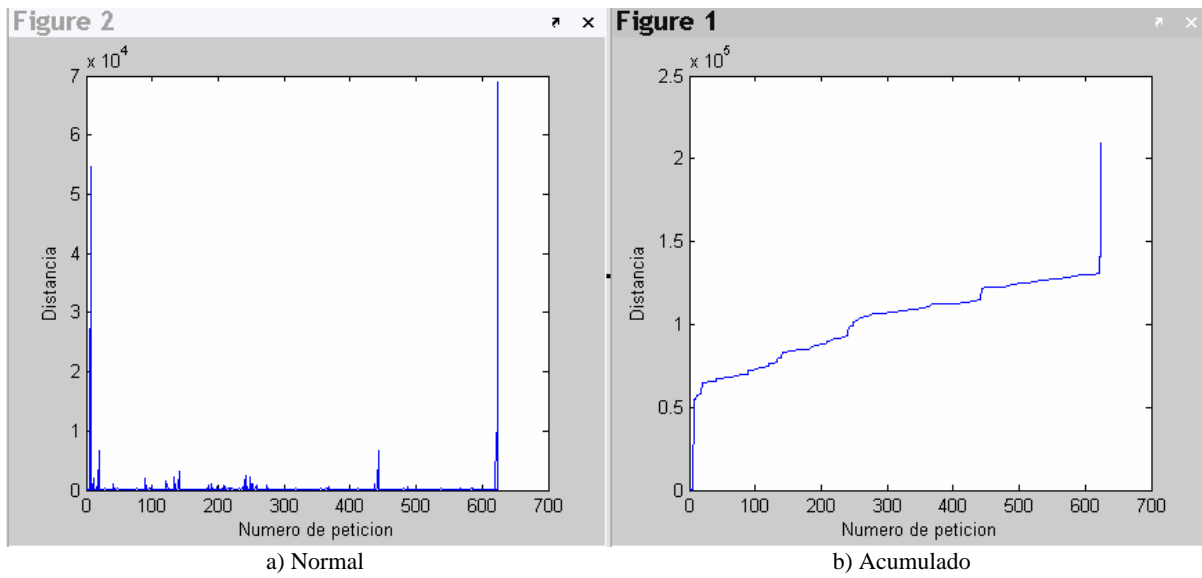


Figura 4.9. Distancias entre peticiones a 3 para el día 24/03/04

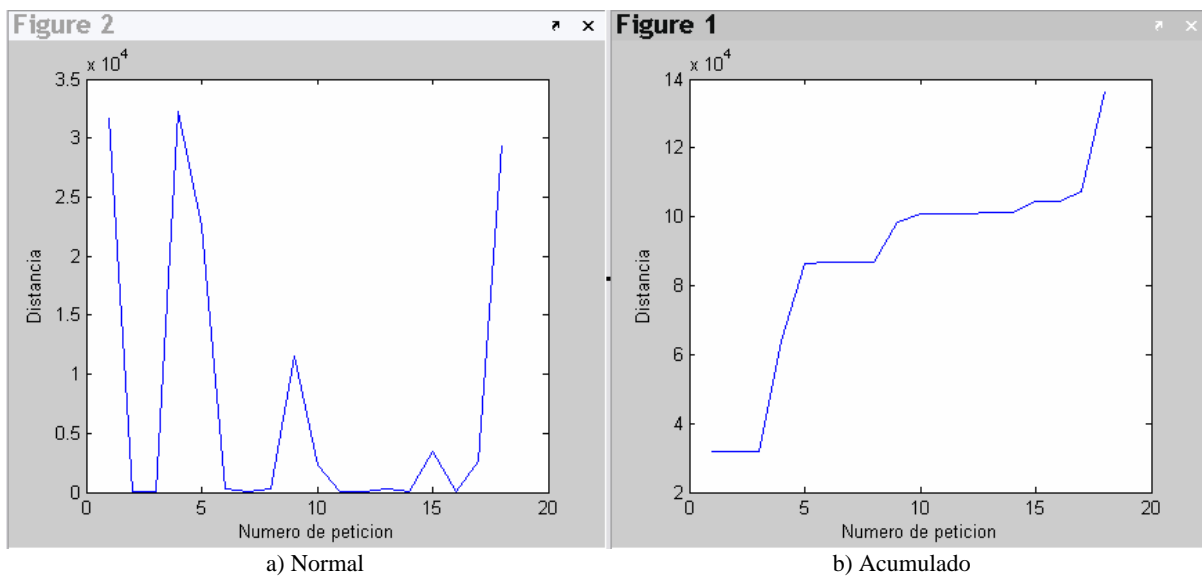


Figura 4.10. Distancias entre peticiones a 3 para el día 25/03/04

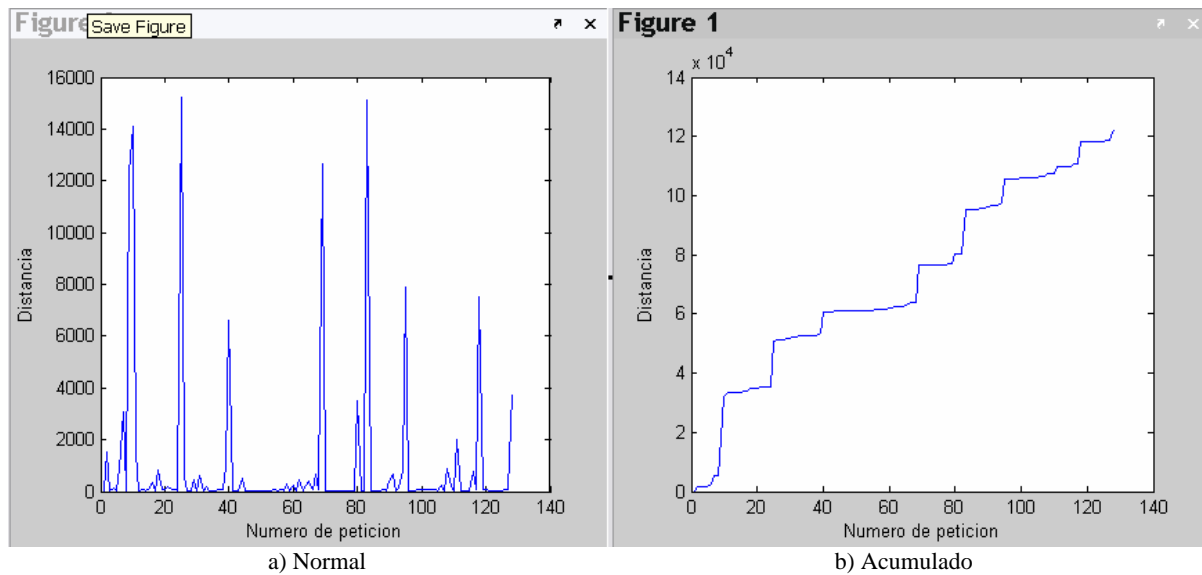


Figura 4.11. Distancias entre peticiones a 3 para el día 26/03/04

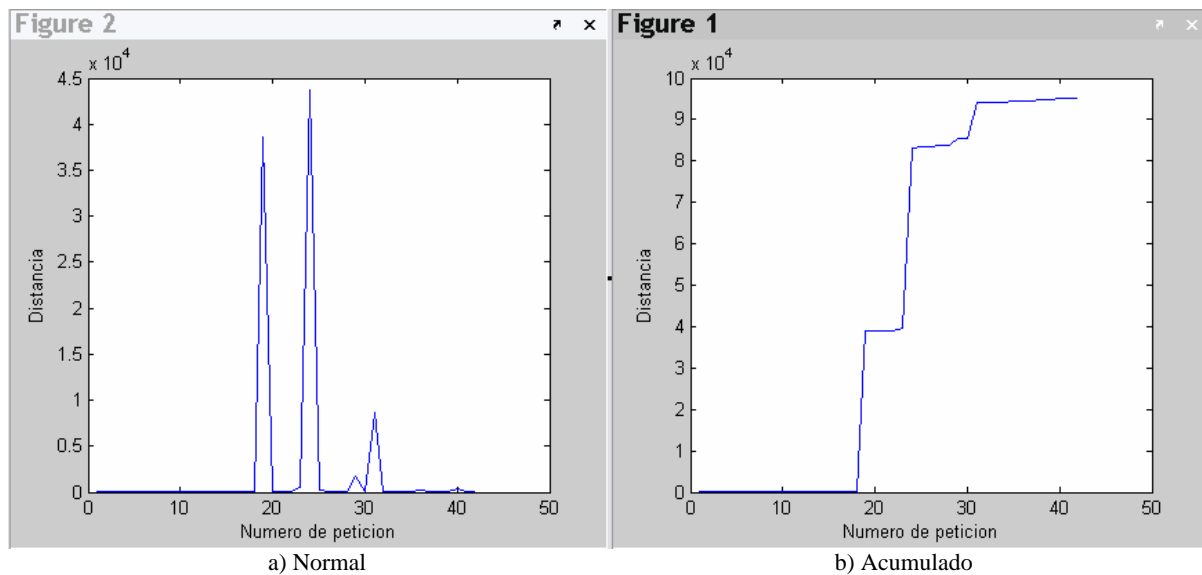


Figura 4.12. Distancias entre peticiones a 3 para el día 27/03/04

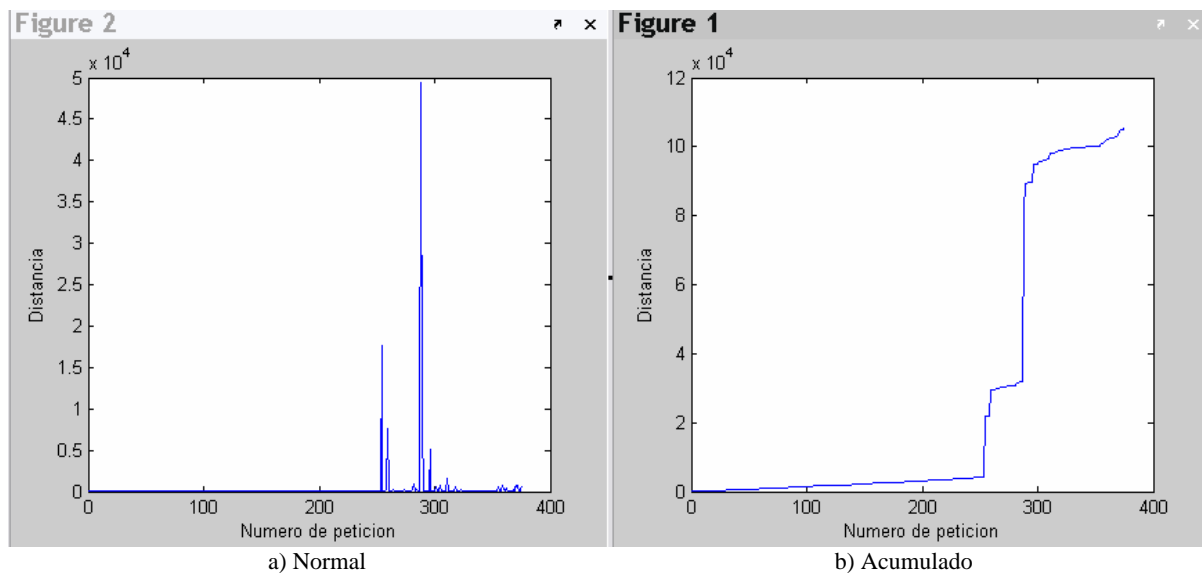


Figura 4.13. Distancias entre peticiones a 3 para el día 28/03/04

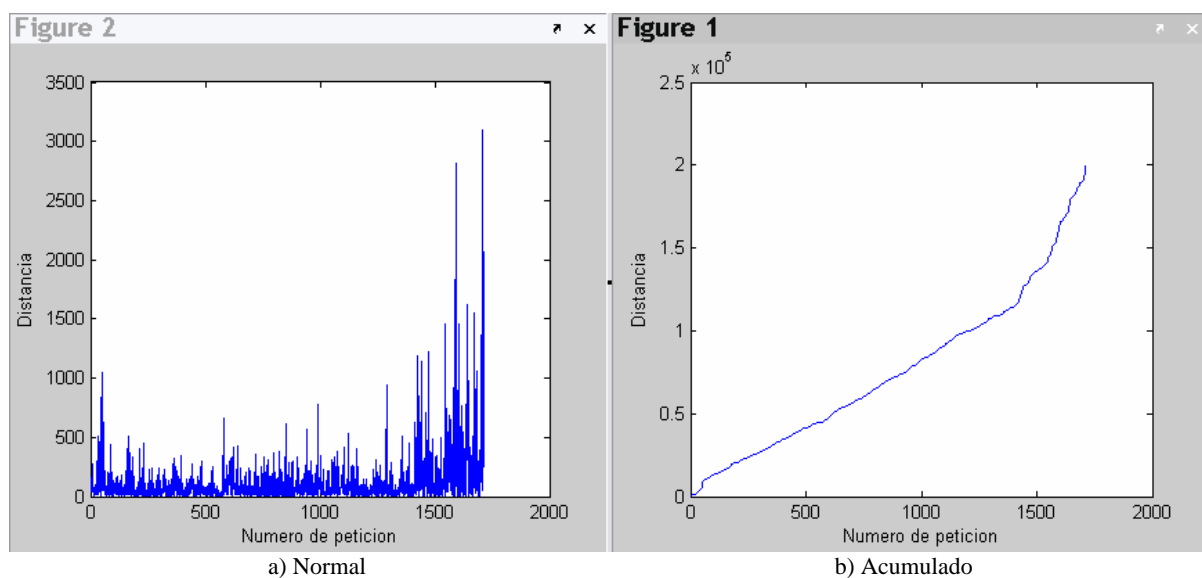


Figura 4.14. Distancias entre peticiones a 4 para el día 22/03/04

A pesar de haber utilizado para el estudio los cuatro objetos más solicitados para el primer día, se observa que solo el objeto 3 es solicitado a lo largo de toda la semana. No es así para los objetos 1 y 2 que pasan de ser solicitados unas 3500 veces el tercer día, al olvido a partir del cuarto. Es también el caso del objeto 4, que es solicitado solo durante el primer día.

Los objetos 1 y 2 son solicitados de forma muy similar lo que hace pensar en algún tipo de dependencia entre ellos. A pesar de fuertes picos en sus gráficas que disparan la distancia con la que se solicitan, estos son escasos y se podría decir que se solicitan de manera homogénea a lo largo de los días, teniendo su media entorno a 3,5.

El caso del objeto 3 es, por el contrario, algo más irregular. A lo largo de los tres primeros días el objeto va cayendo lentamente en desuso (pasando de 171 objetos solicitados entre peticiones para el primer día, a 228 y 337, para segundo y el tercero). En los siguientes cuatro días, la popularidad del objeto 3 decae drásticamente, y sufre altibajos a lo largo de los mismos, tanto en número de veces que se solicita como en espacio entre peticiones.

Por último en el objeto 4 puede observarse una gráfica bastante regular (con media de 117), y que en su último tramo desvela como el objeto va cayendo en el desuso al finalizar el día.

4.1.2. TAMAÑOS ENTRE PETICIONES

Las graficas que se exponen a continuación muestran los tamaños existentes entre sucesivas peticiones a cada uno de los cuatro objetos anteriormente citados, esto es, representan el número de bytes solicitados entre sucesivas peticiones a un objeto dado respecto al tiempo.

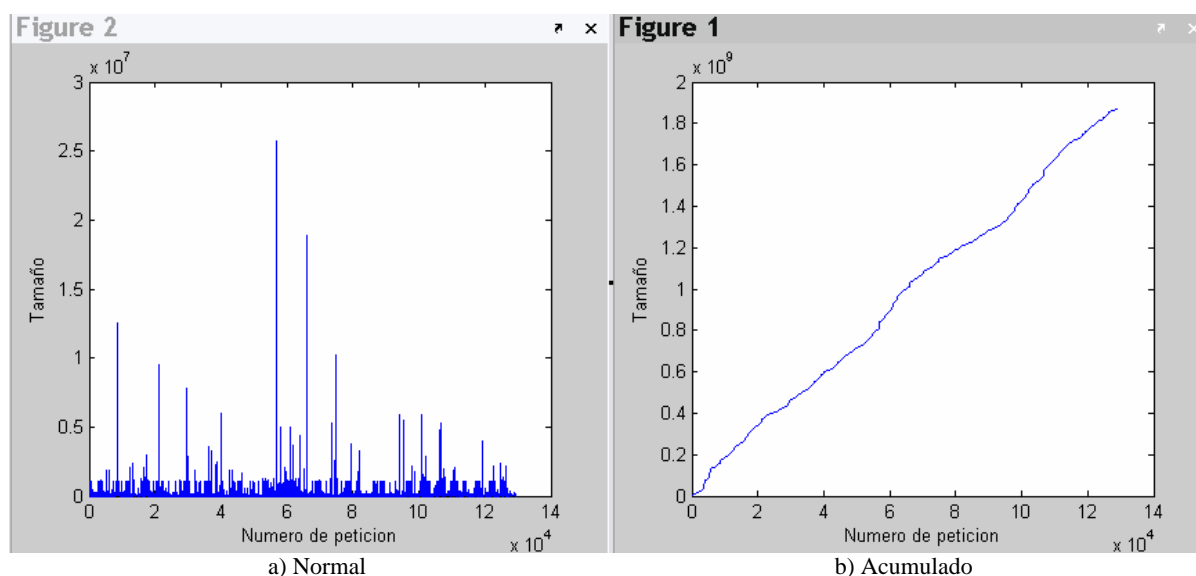


Figura 4.15. Tamaños entre peticiones a 1 para el día 22/03/04

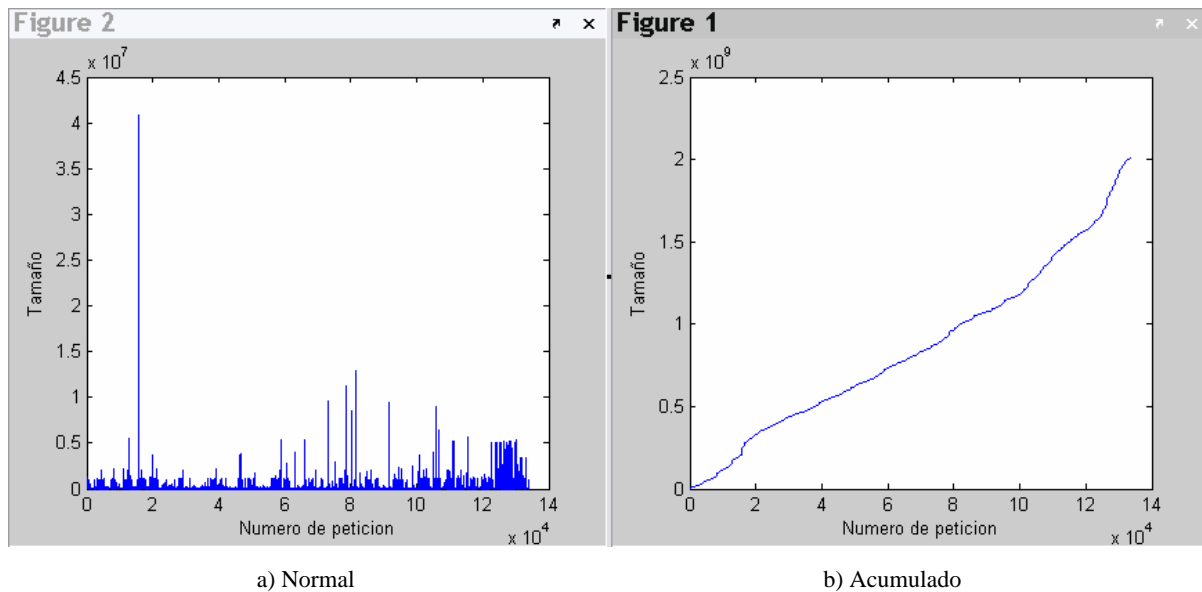


Figura 4.16. Tamaños entre peticiones a 1 para el día 23/03/04

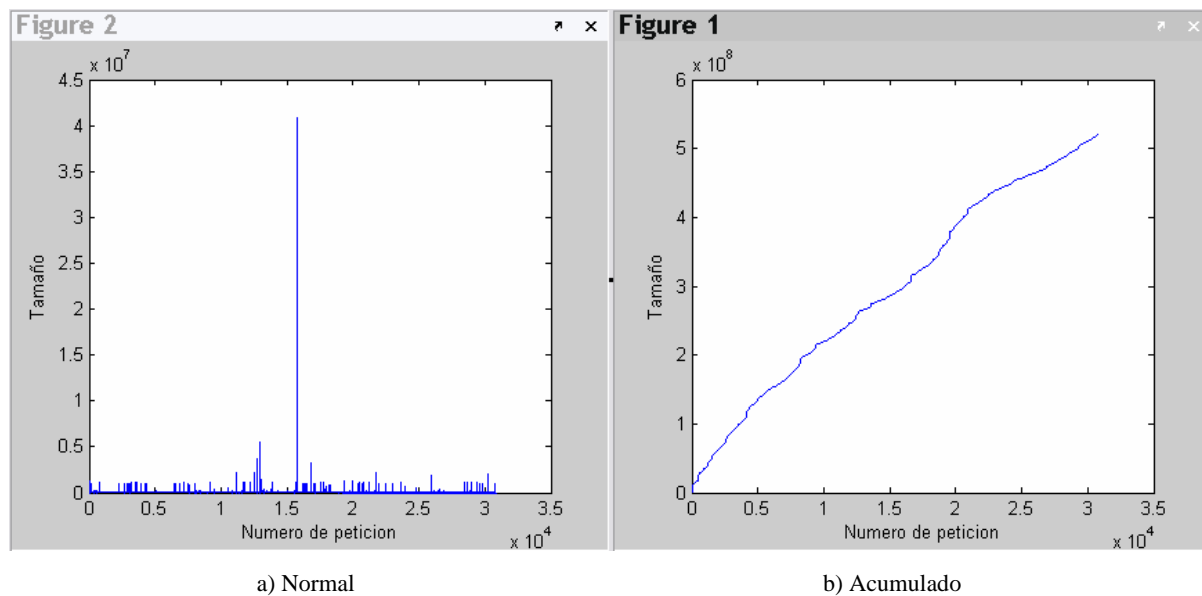


Figura 4.17. Tamaños entre peticiones a 1 para el día 24/03/04

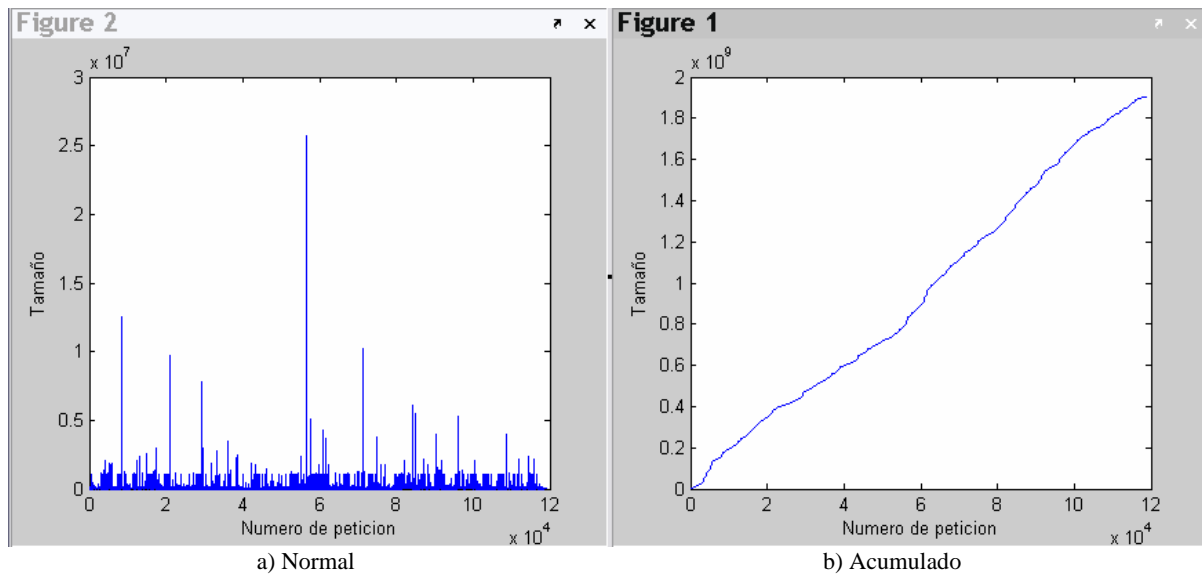


Figura 4.18. Tamaños entre peticiones a 2 para el día 22/03/04

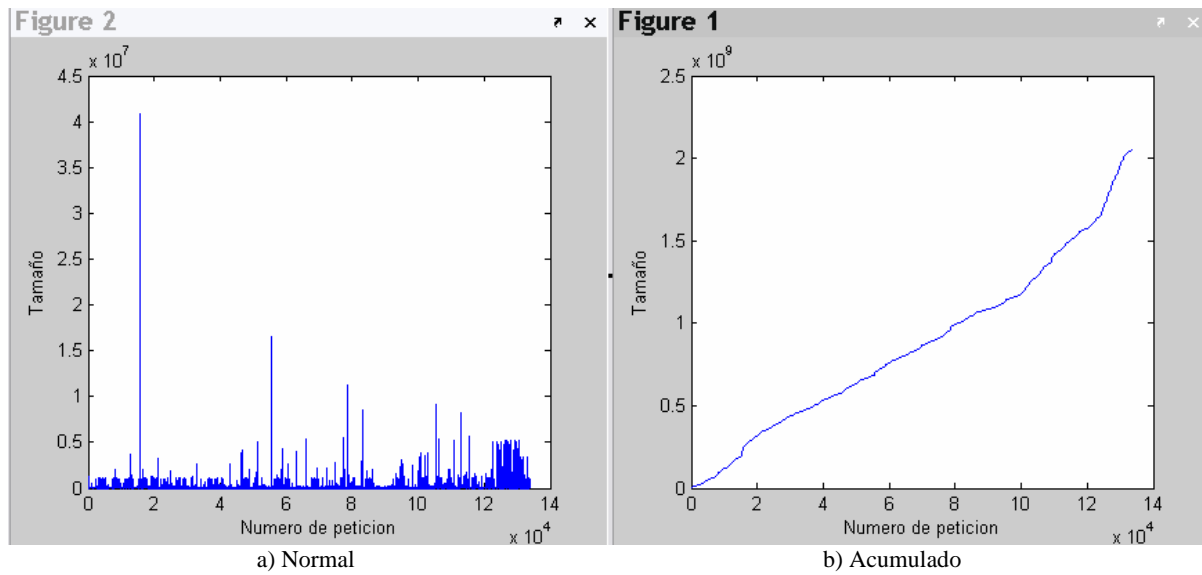


Figura 4.19. Tamaños entre peticiones a 2 para el día 23/03/04

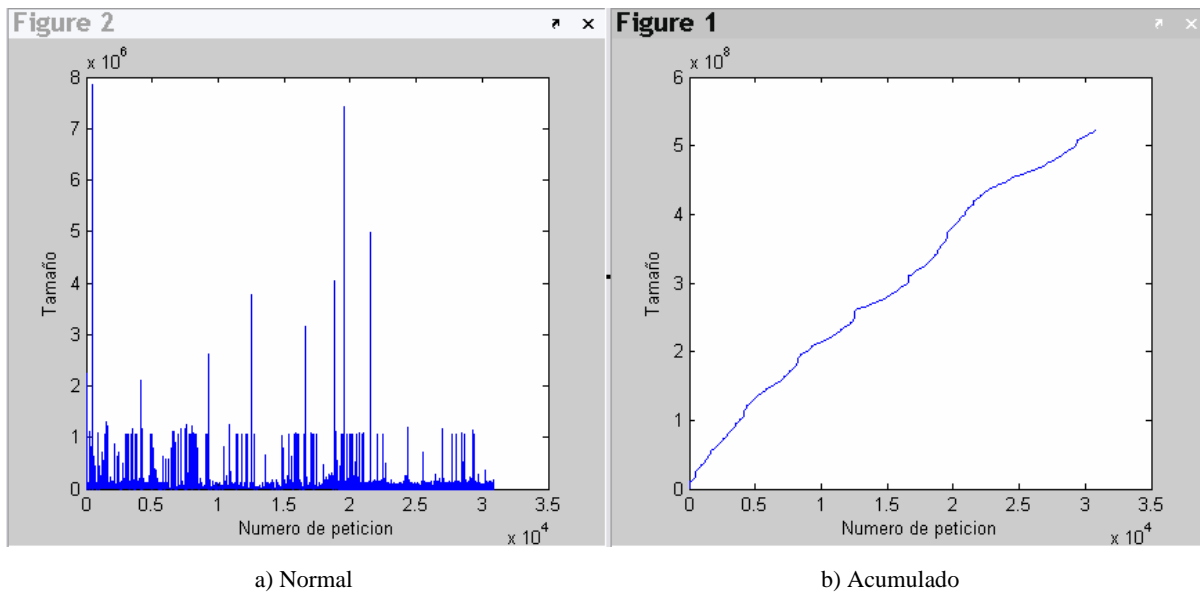


Figura 4.20. Tamaños entre peticiones a 2 para el día 24/03/04

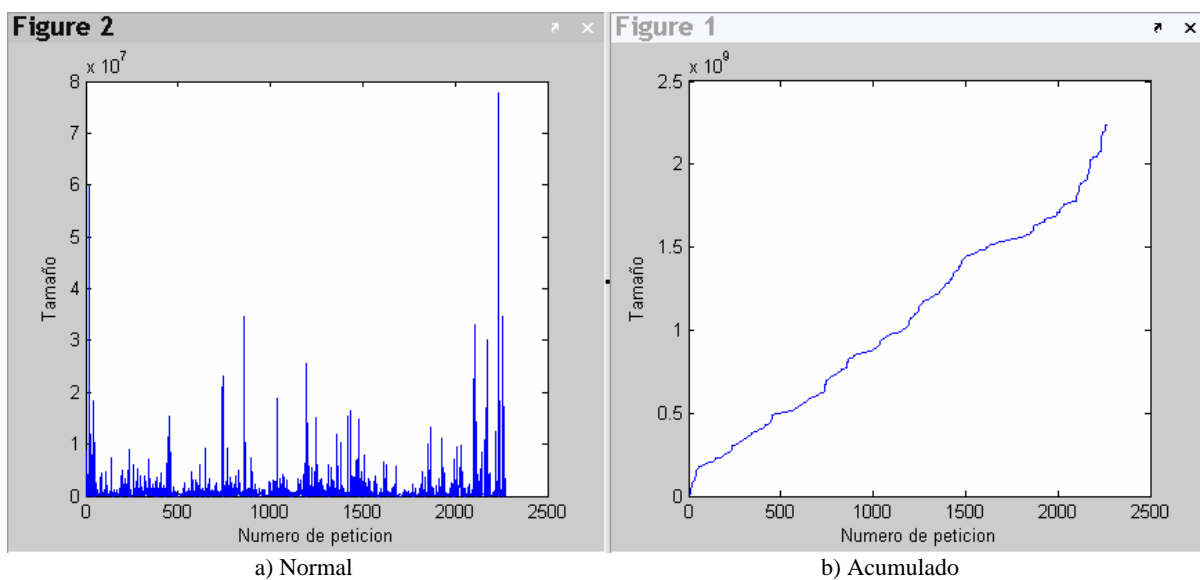


Figura 4.21. Tamaños entre peticiones a 3 para el día 22/03/04

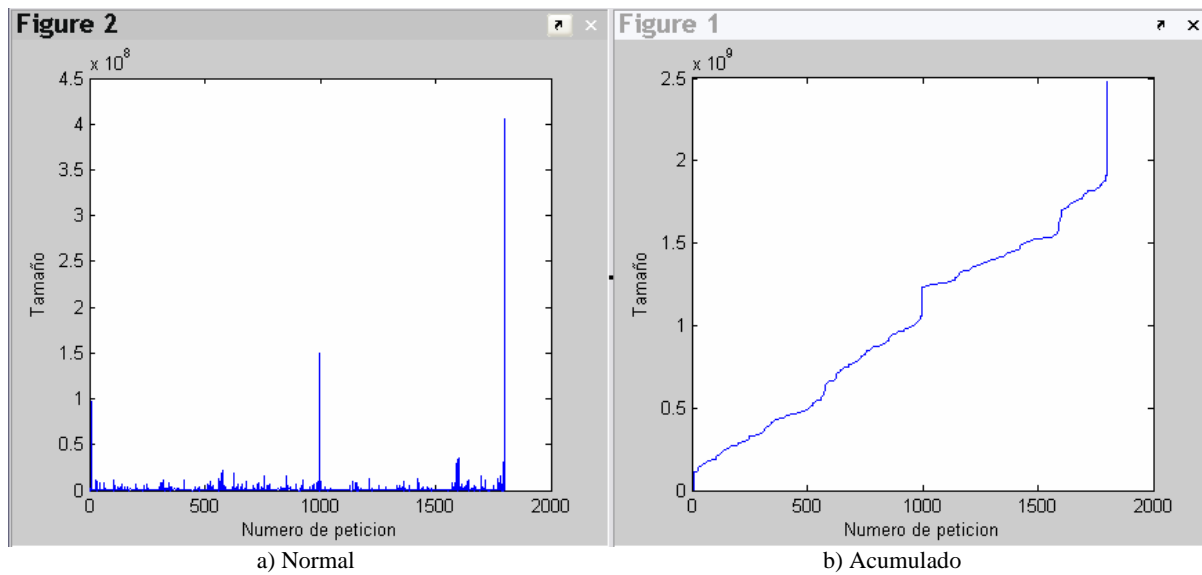


Figura 4.22. Tamaños entre peticiones a 3 para el día 23/03/04

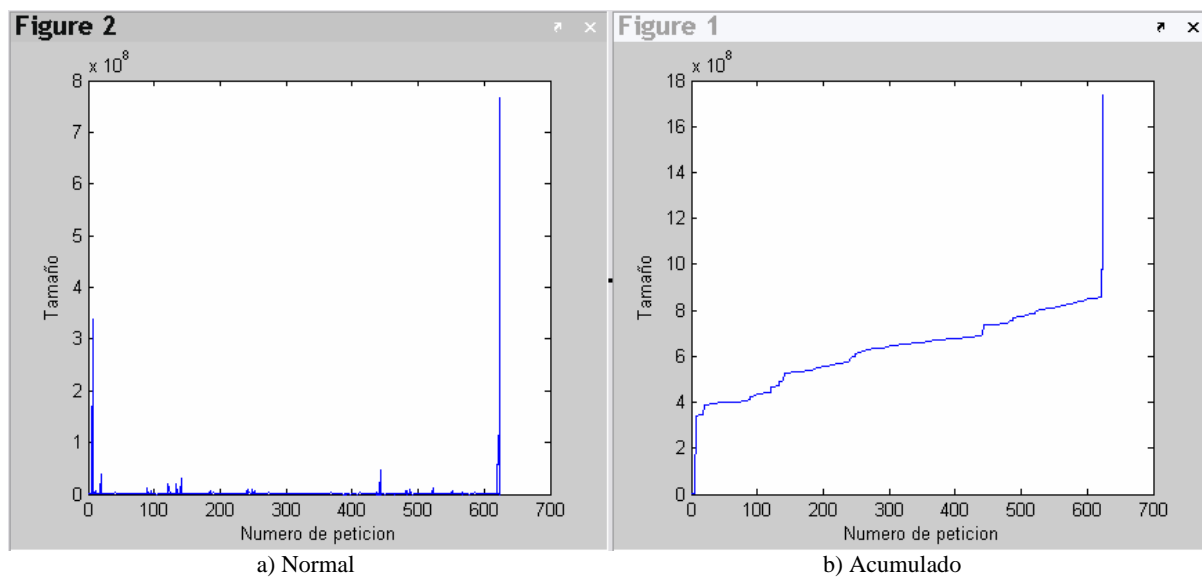


Figura 4.23. Tamaños entre peticiones a 3 para el día 24/03/04

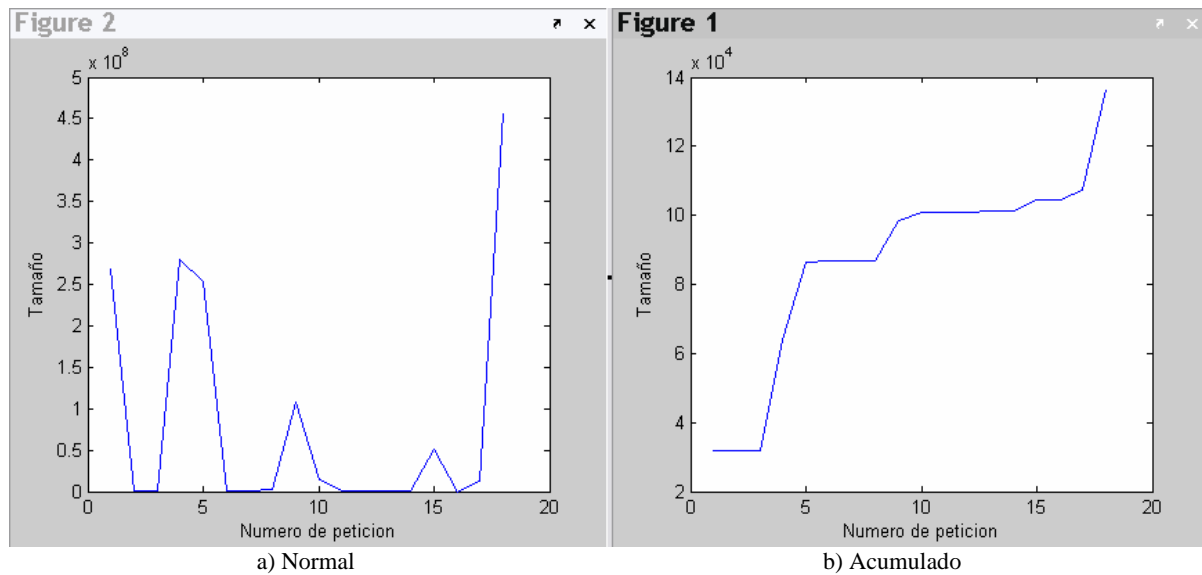


Figura 4.24. Tamaños entre peticiones a 3 para el día 25/03/04

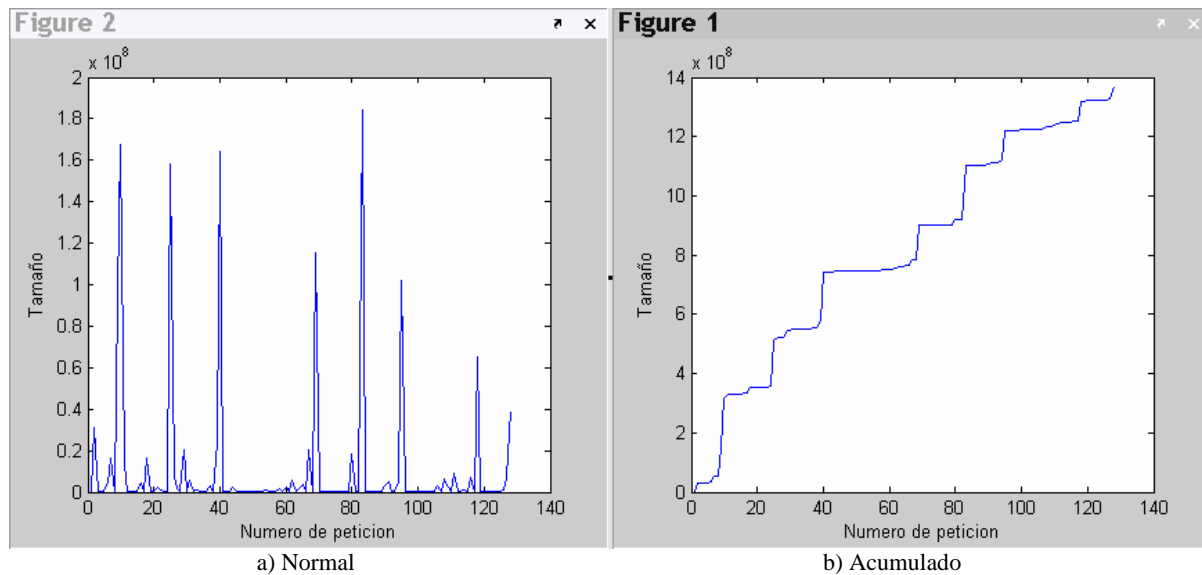


Figura 4.25. Tamaños entre peticiones a 3 para el día 26/03/04

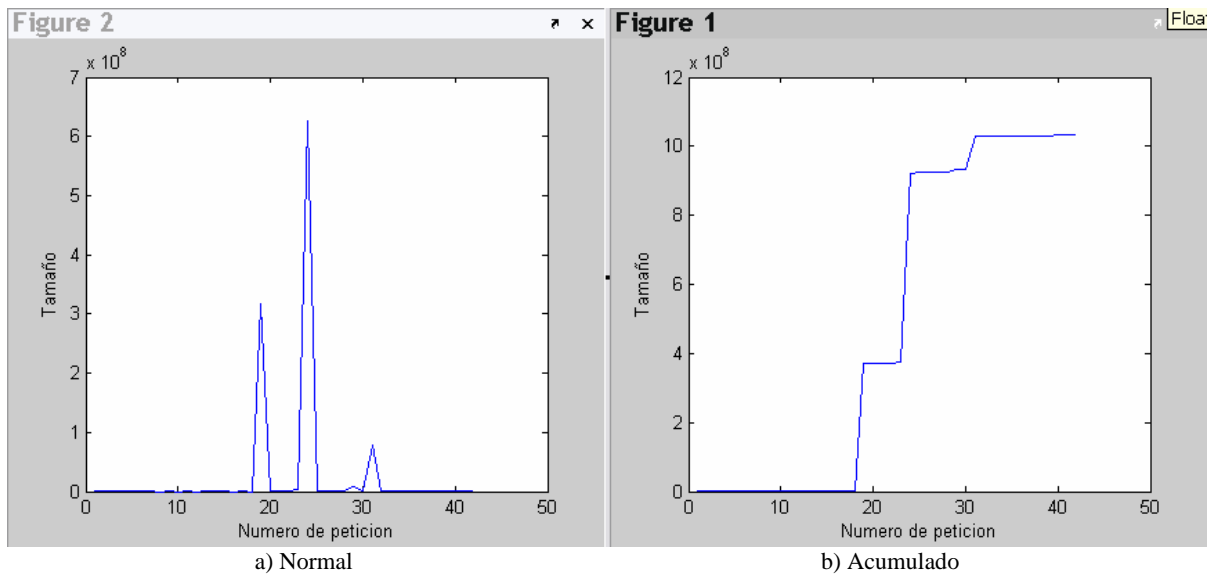


Figura 4.26. Tamaños entre peticiones a 3 para el día 27/03/04

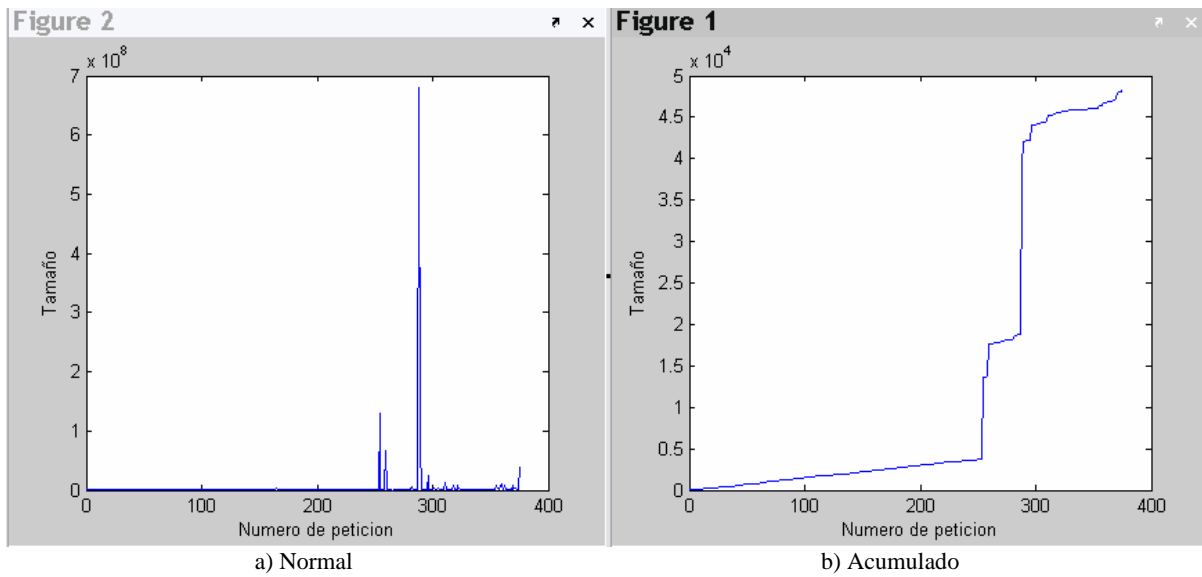


Figura 4.27. Tamaños entre peticiones a 3 para el día 28/03/04

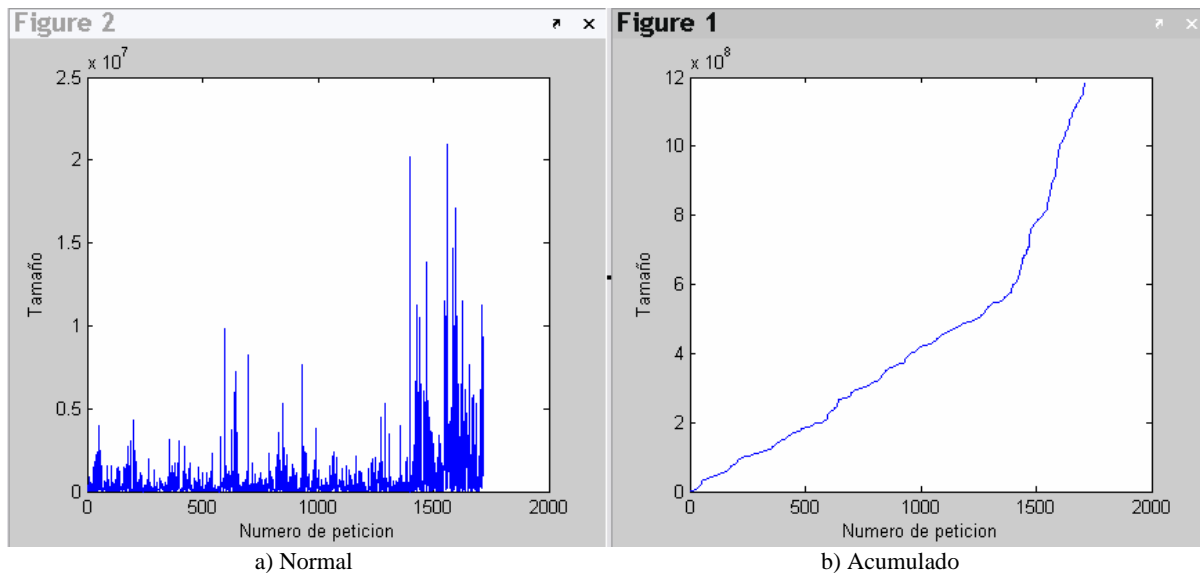


Figura 4.28. Tamaños entre peticiones a 4 para el día 22/03/04

Los objetos 1 y 2 presentan gráficas prácticamente idénticas, y a pesar de fuertes picos en sus gráficas que disparan los tamaños entre peticiones, estos son escasos y se podría decir que se solicitan de manera homogénea a lo largo de los días, teniendo su media entorno a 15 KB.

El caso del objeto 3 es, por el contrario, algo más irregular. A lo largo de los tres primeros días el objeto va cayendo lentamente en desuso (pasando de 980 KB de tamaño entre peticiones para el primer día, a 1300 KB y 1600 KB, para segundo y el tercero). En los siguientes cuatro días, la popularidad del objeto 3 decae drásticamente, y sufre altibajos a lo largo de los mismos, tanto en número de veces que se solicita como en tamaño entre solicitudes.

Por último en el objeto 4 puede observarse una gráfica bastante regular (con una media de tamaño entre peticiones de 13 KB), y que en su último tramo desvela como el objeto va cayendo en el desuso al finalizar el día.

4.1.3. TIEMPOS ENTRE PETICIONES

Las gráficas que se exponen a continuación muestran el tiempo existente entre sucesivas peticiones a cada uno de los cuatro objetos anteriormente citados, esto es, representan el número de segundos transcurridos entre sucesivas peticiones a un objeto dado respecto al tiempo.

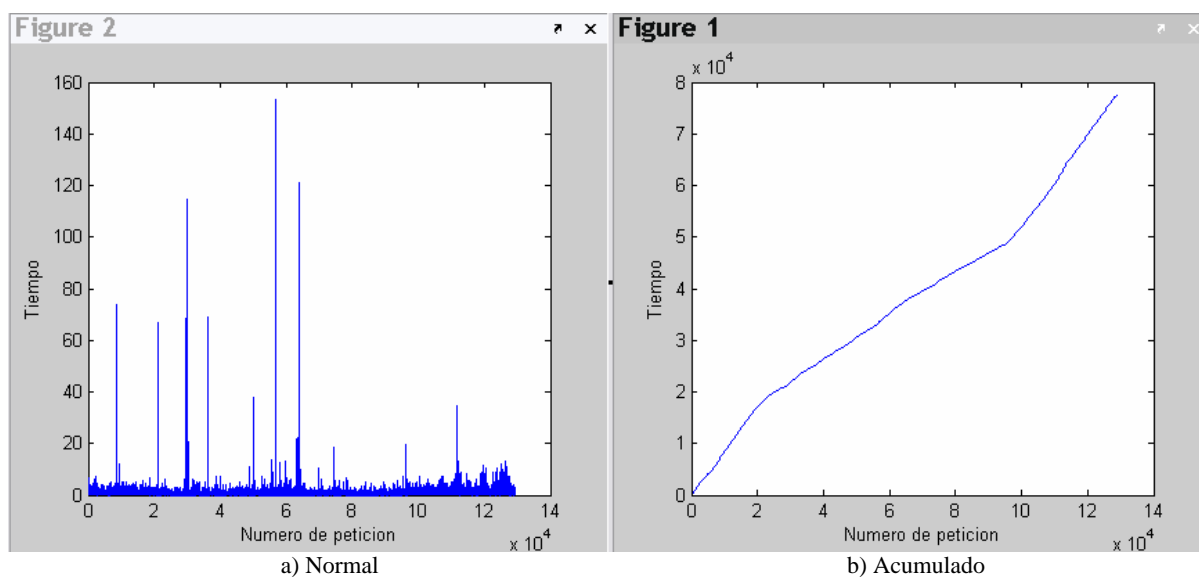


Figura 4.29. Tiempos entre peticiones a 1 para el día 22/03/04

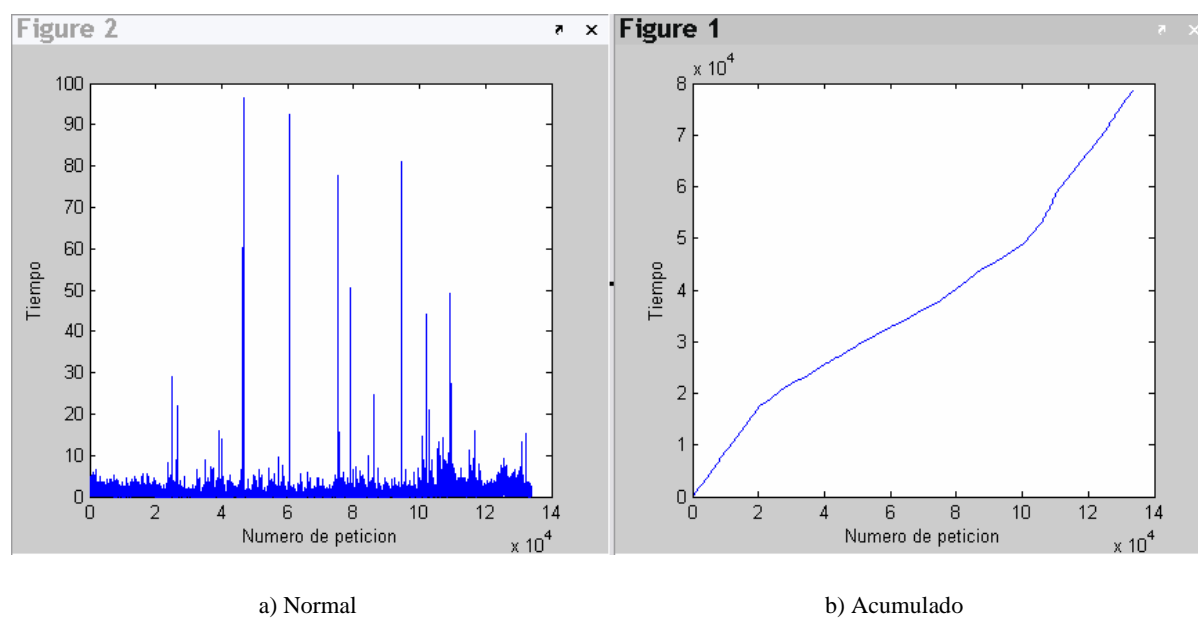
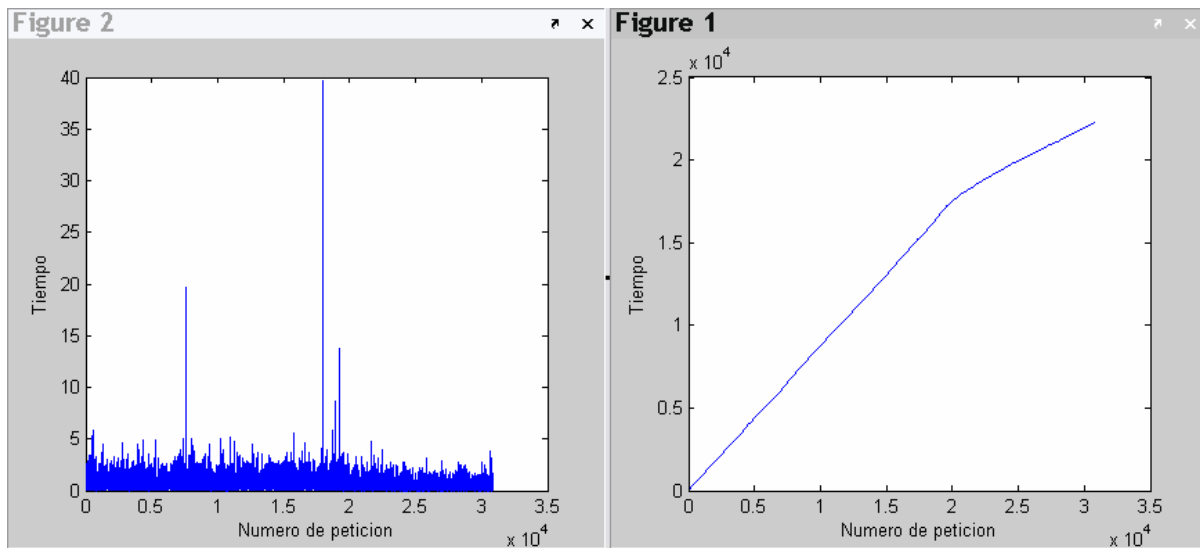


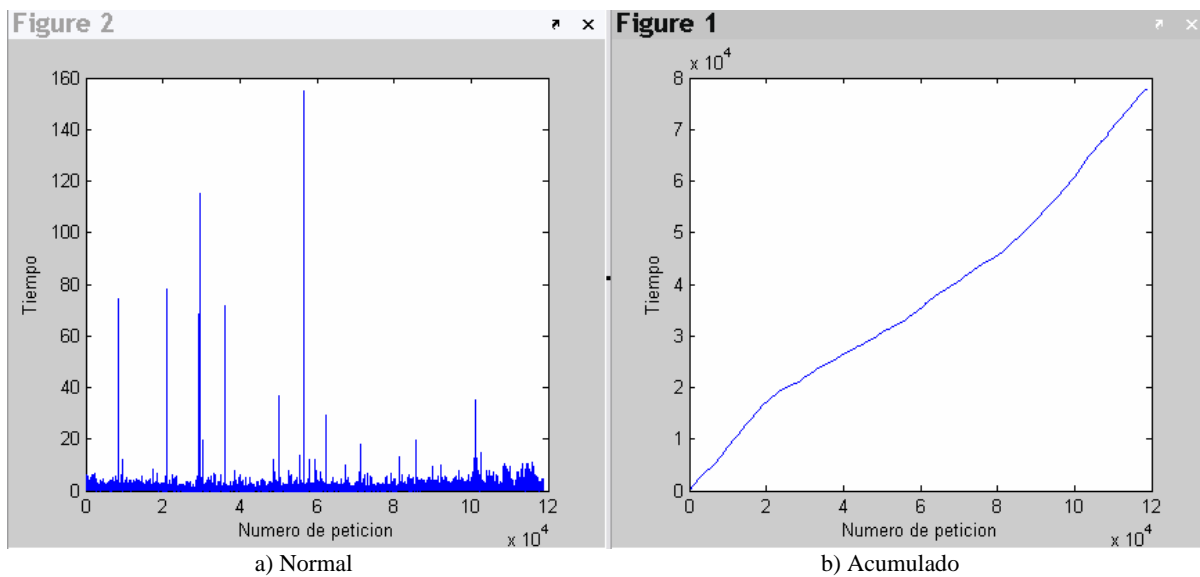
Figura 4.30. Tiempos entre peticiones a 1 para el día 23/03/04



a) Normal

b) Acumulado

Figura 4.31. Tiempos entre peticiones a 1 para el día 24/03/04



a) Normal

b) Acumulado

Figura 4.32. Tiempos entre peticiones a 2 para el día 22/03/04

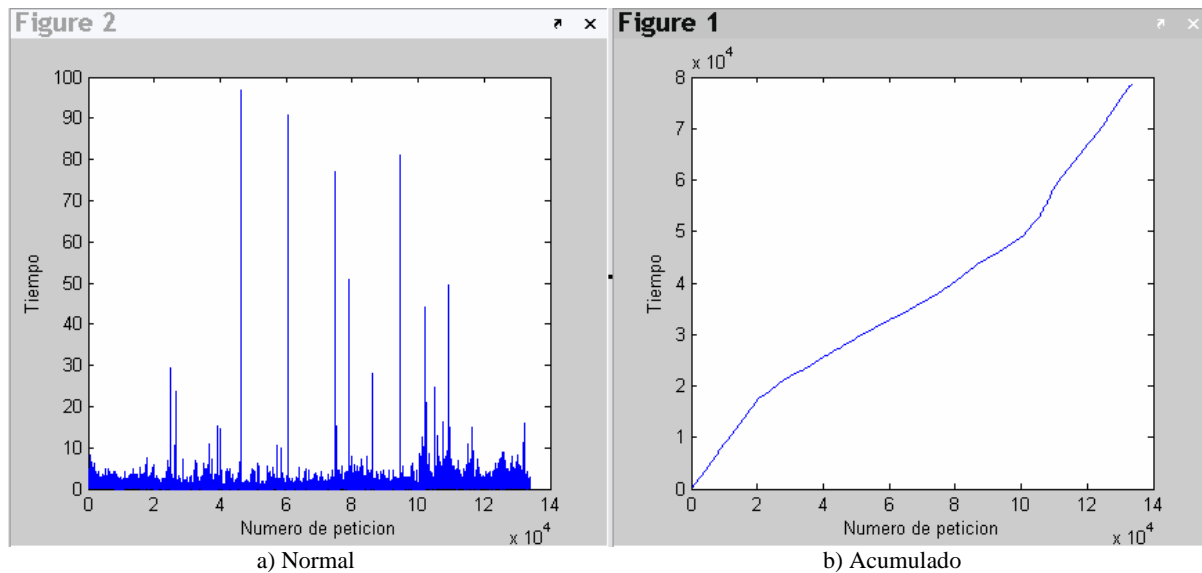


Figura 4.33. Tiempos entre peticiones a 2 para el día 23/03/04

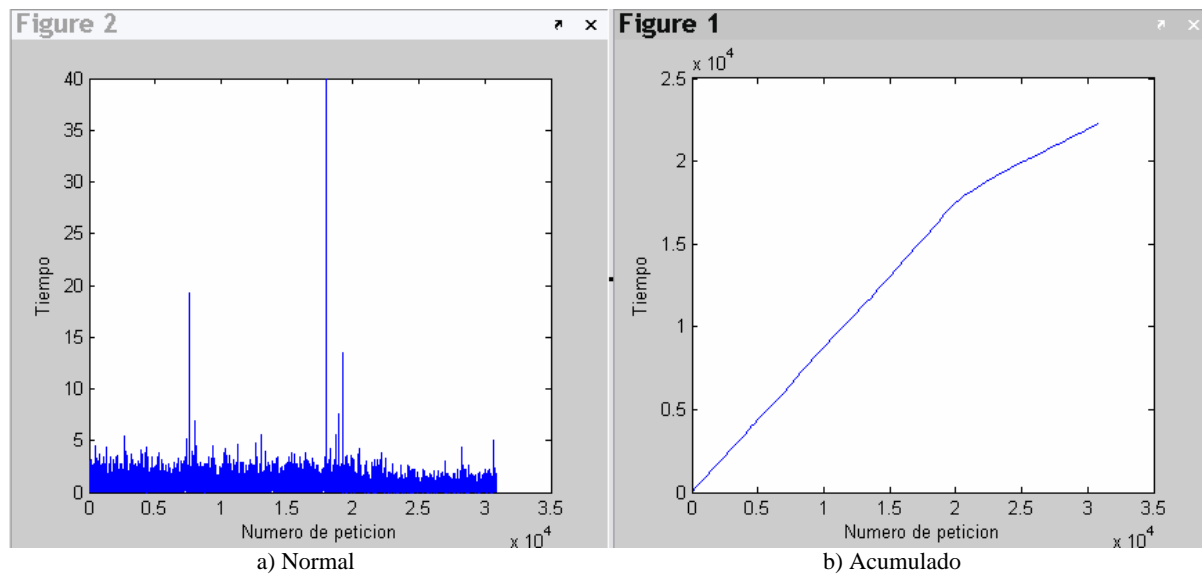


Figura 4.34. Tiempos entre peticiones a 2 para el día 24/03/04

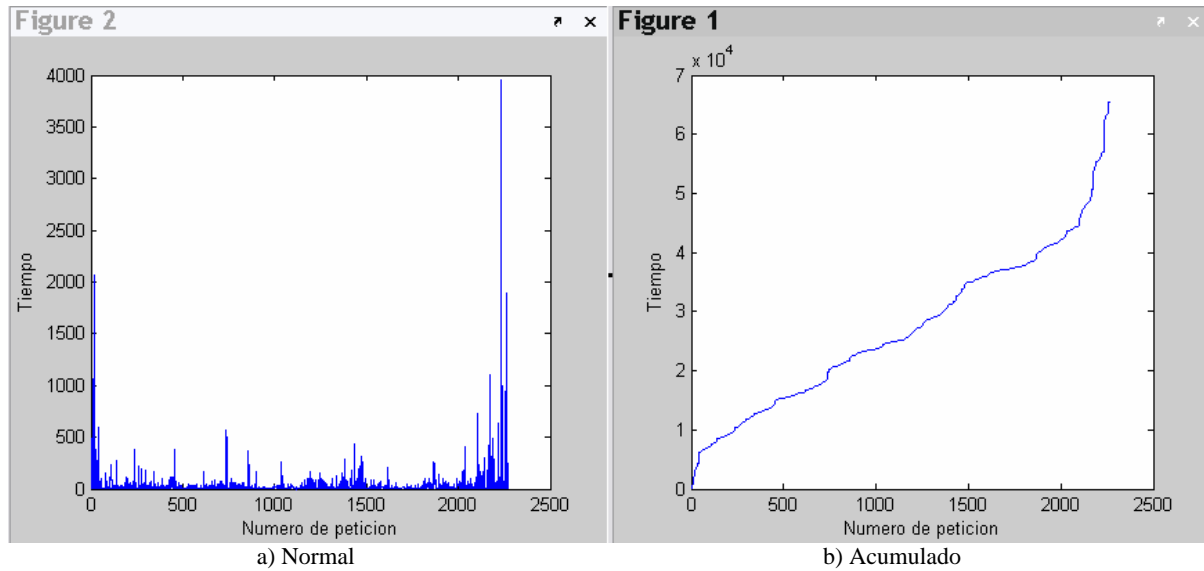


Figura 4.35. Tiempos entre peticiones a 3 para el día 22/03/04

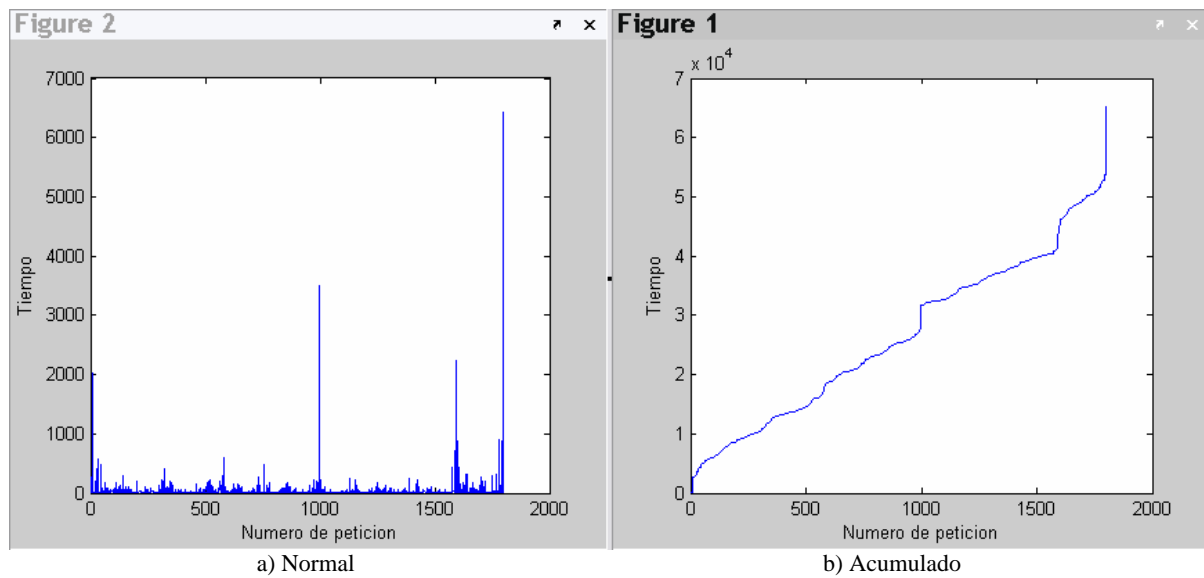


Figura 4.36. Tiempos entre peticiones a 3 para el día 23/03/04

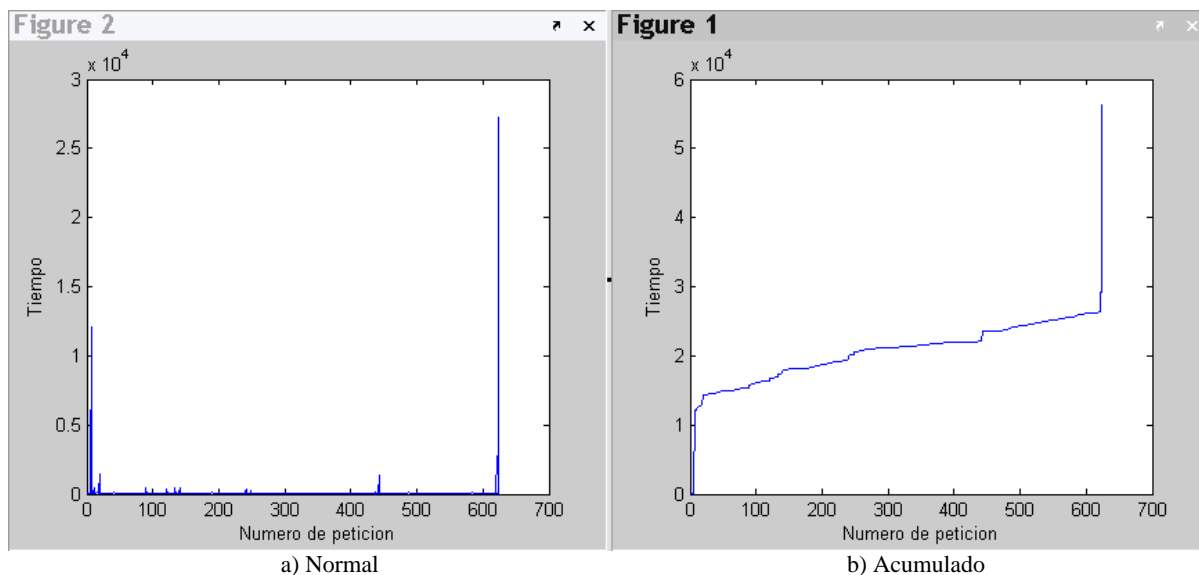


Figura 4.37. Tiempos entre peticiones a 3 para el día 24/03/04

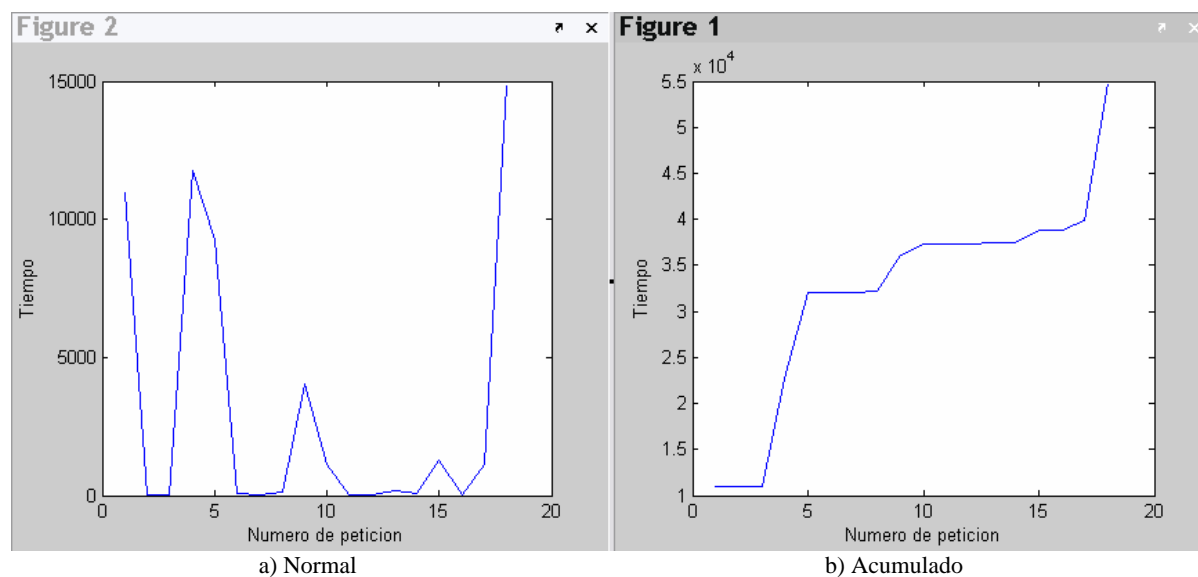


Figura 4.38. Tiempos entre peticiones a 3 para el día 25/03/04

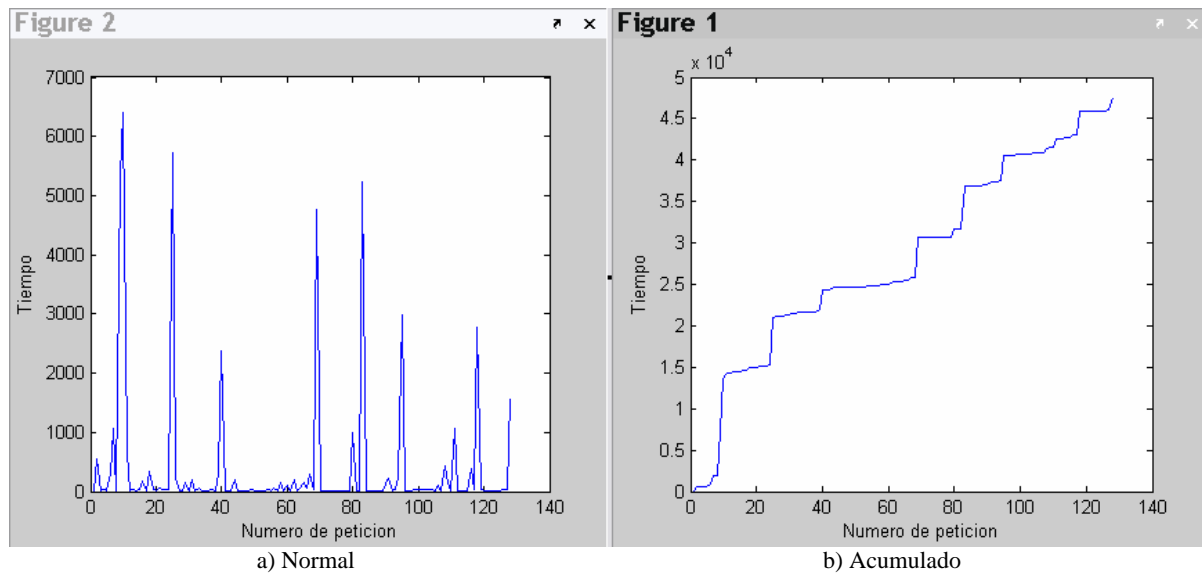


Figura 4.39. Tiempos entre peticiones a 3 para el día 26/03/04

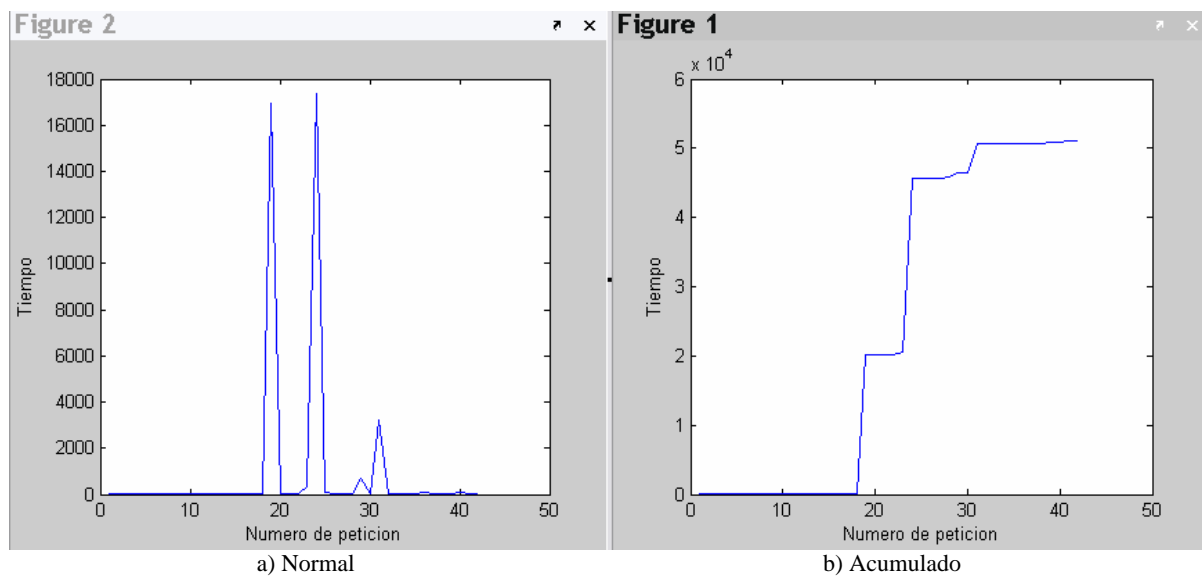


Figura 4.40. Tiempos entre peticiones a 3 para el día 27/03/04

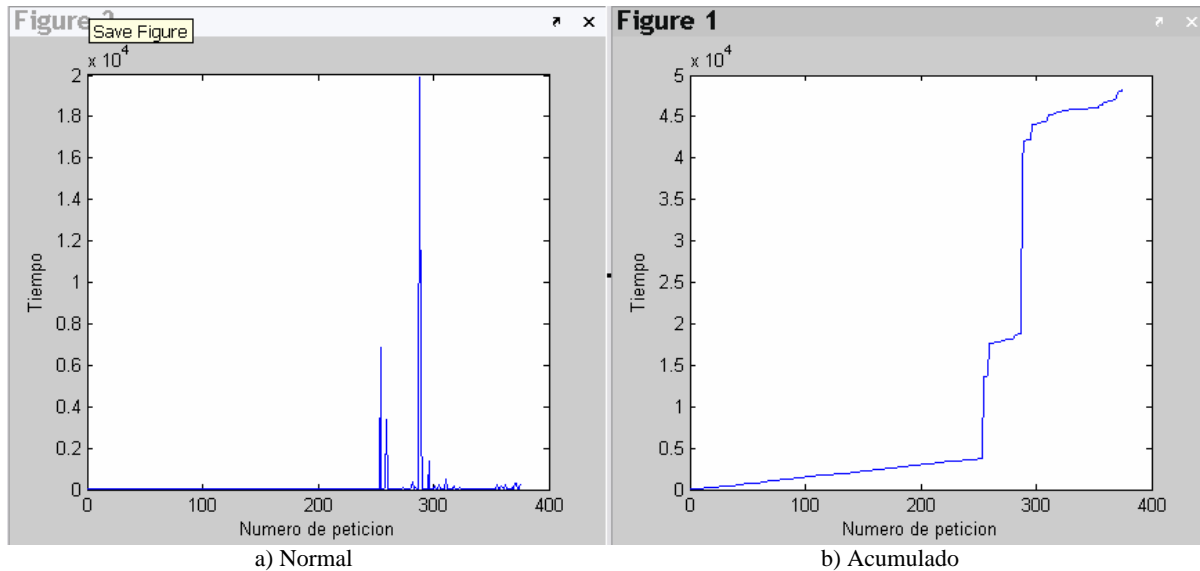


Figura 4.41. Tiempos entre peticiones a 3 para el día 28/03/04

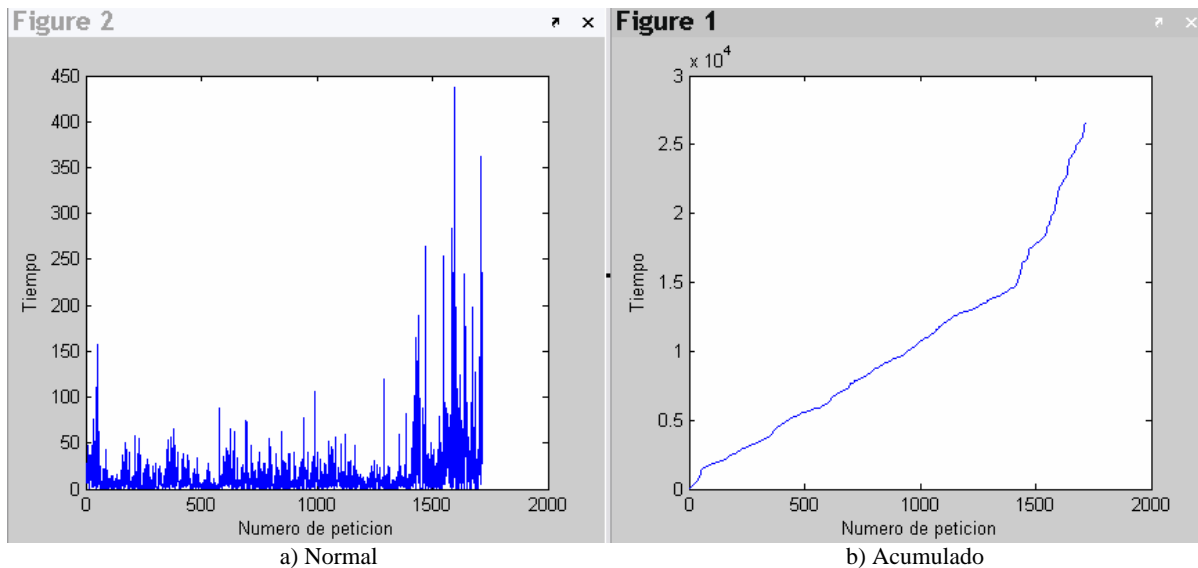


Figura 4.42. Tiempos entre peticiones a 4 para el día 22/03/04

Los objetos 1 y 2 son solicitados de forma muy similar lo que hace pensar en algún tipo de dependencia entre ellos. A pesar de fuertes picos en sus gráficas que disparan el tiempo entre peticiones, estos son escasos y se podría decir que se solicitan de manera homogénea a lo largo de los días, teniendo su media entorno a 0.65 sg.

El caso del objeto 3 es, por el contrario, algo más irregular. A lo largo de los tres primeros días el objeto va cayendo lentamente en desuso (pasando de 29 segundos de tiempo entre peticiones para el primer día, a 36 y 90, para segundo y el tercero). En los siguientes cuatro días, la popularidad del objeto 3 decae drásticamente, y sufre altibajos a lo largo de los mismos, tanto en número de veces que se solicita como en tamaño entre solicitudes.

Por último en el objeto 4 puede observarse una gráfica bastante regular (con media de 15 sg), y que en su último tramo desvela como el objeto va cayendo en el desuso al finalizar el día.

3.3.4. CONCLUSIONES

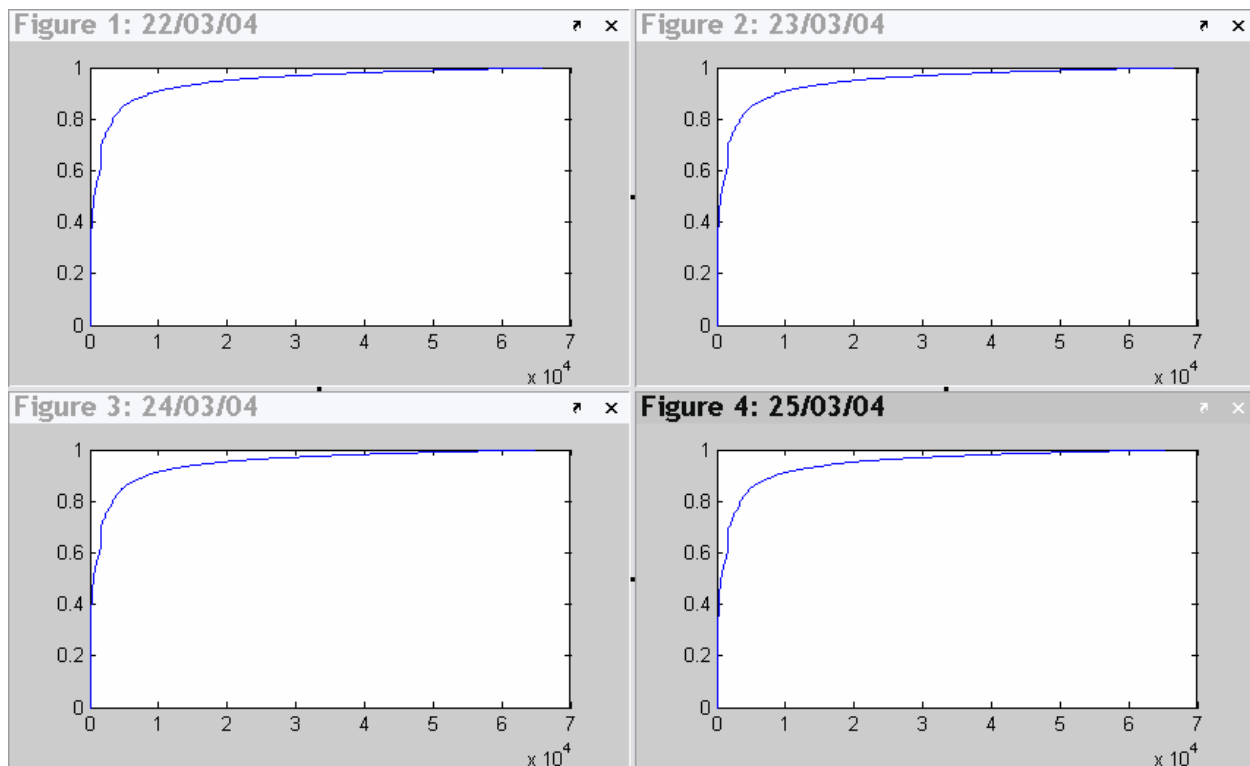
En este estudio cabe destacar la breve popularidad de los objetos. Así, los objetos 1 y 2 solo eran referenciados los primeros cuatro días, y el objeto 4 solo el primero. Solo el objeto 3 permanecía a lo largo de la semana de estudio, pero este comenzó a caer en desuso a partir del cuarto día.

4.2. CURVAS DE DISTRIBUCIÓN DE LOS TAMAÑOS

Haciendo uso de la opción ‘Curva CDF de los tamaños’ de la aplicación principal que se expuso en el [apartado 2.5.2.12](#), dedicamos este apartado a la obtención de las curvas de distribución de los tamaños acumulados. Estas representan la probabilidad acumulada ‘Y’ de encontrar un objeto de tamaño ‘X’ bytes.

Se exponen a continuación las gráficas obtenidas del estudio de las trazas del 22 de Marzo de 2004 al 28 de Marzo de 2004 para ambos servidores proxy (RTP y PA).

4.2.1. SERVIDOR RTP



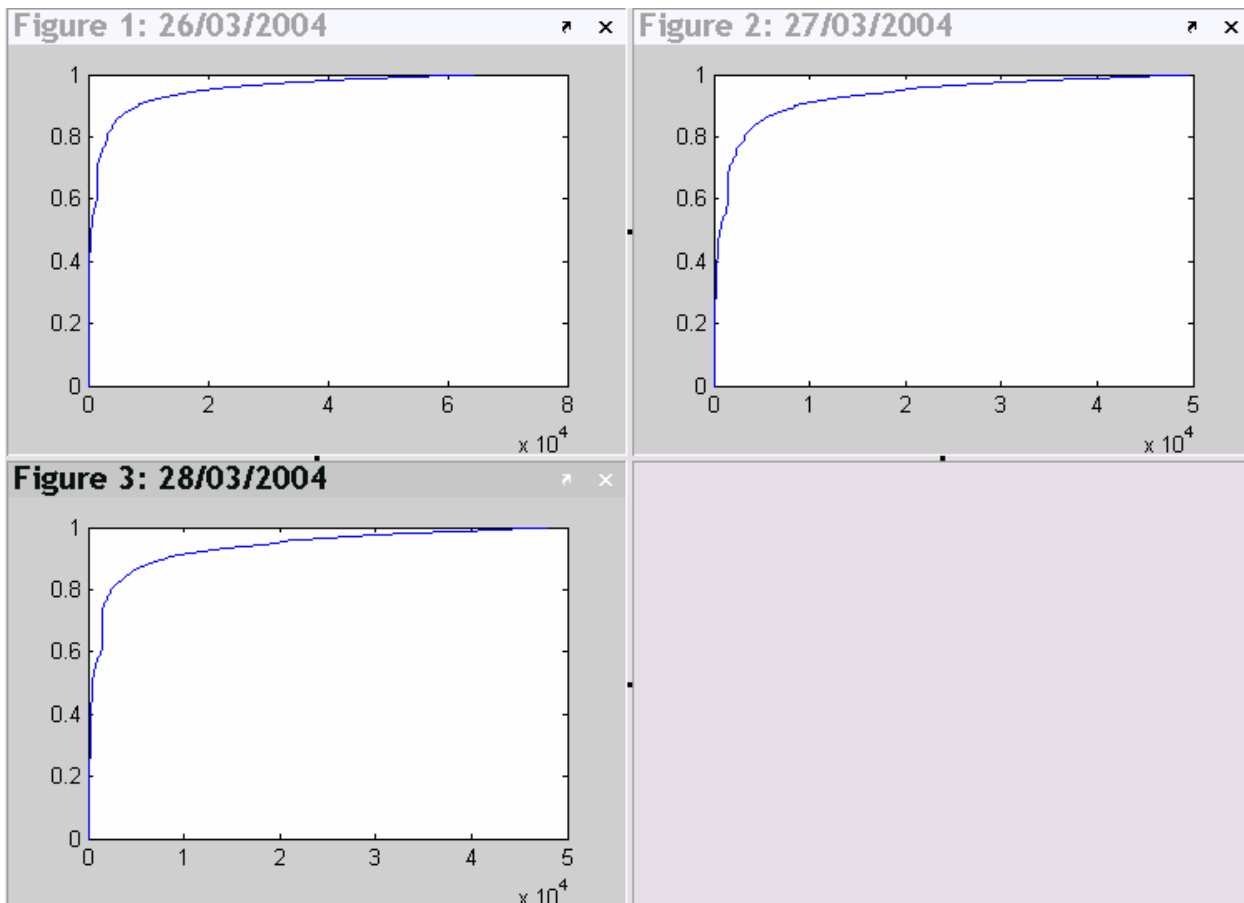


Figura 4.43. Curvas de distribución de los tamaños para el servidor RTP

4.2.2. SERVIDOR PA

Figure 1: 22/03/04

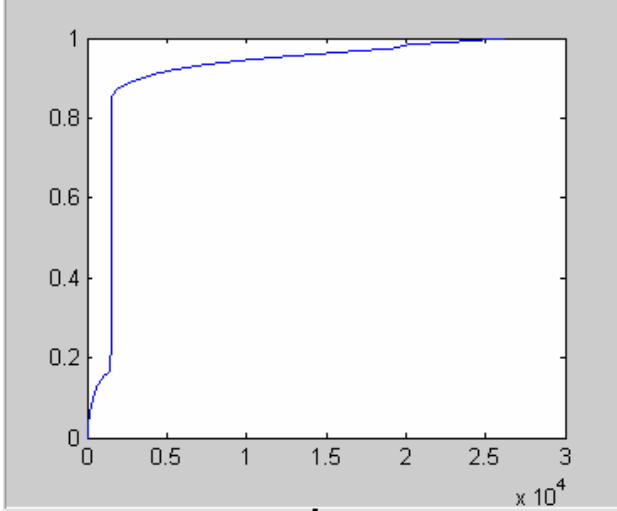


Figure 2: 23/03/04

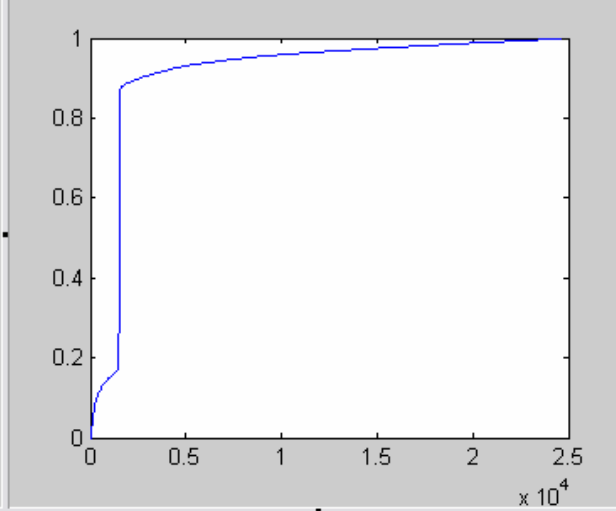


Figure 3: 24/03/04

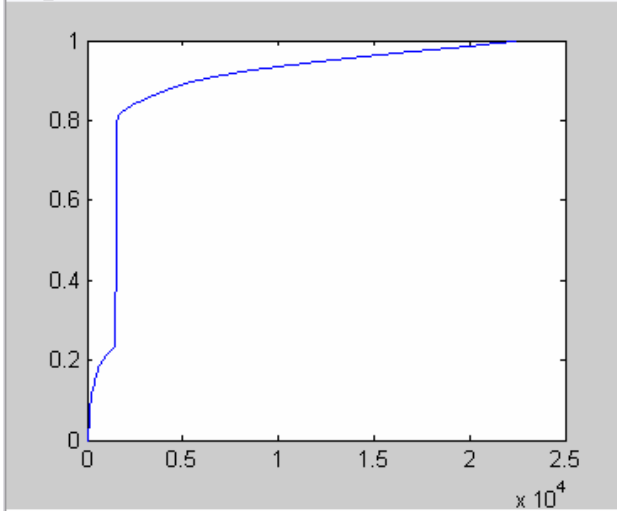
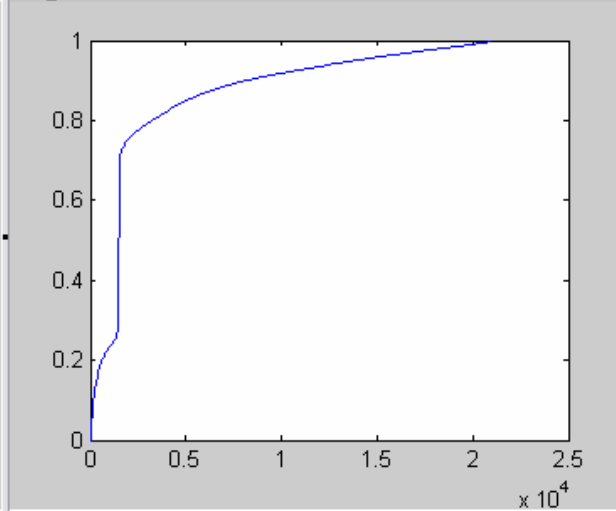


Figure 4: 25/03/04



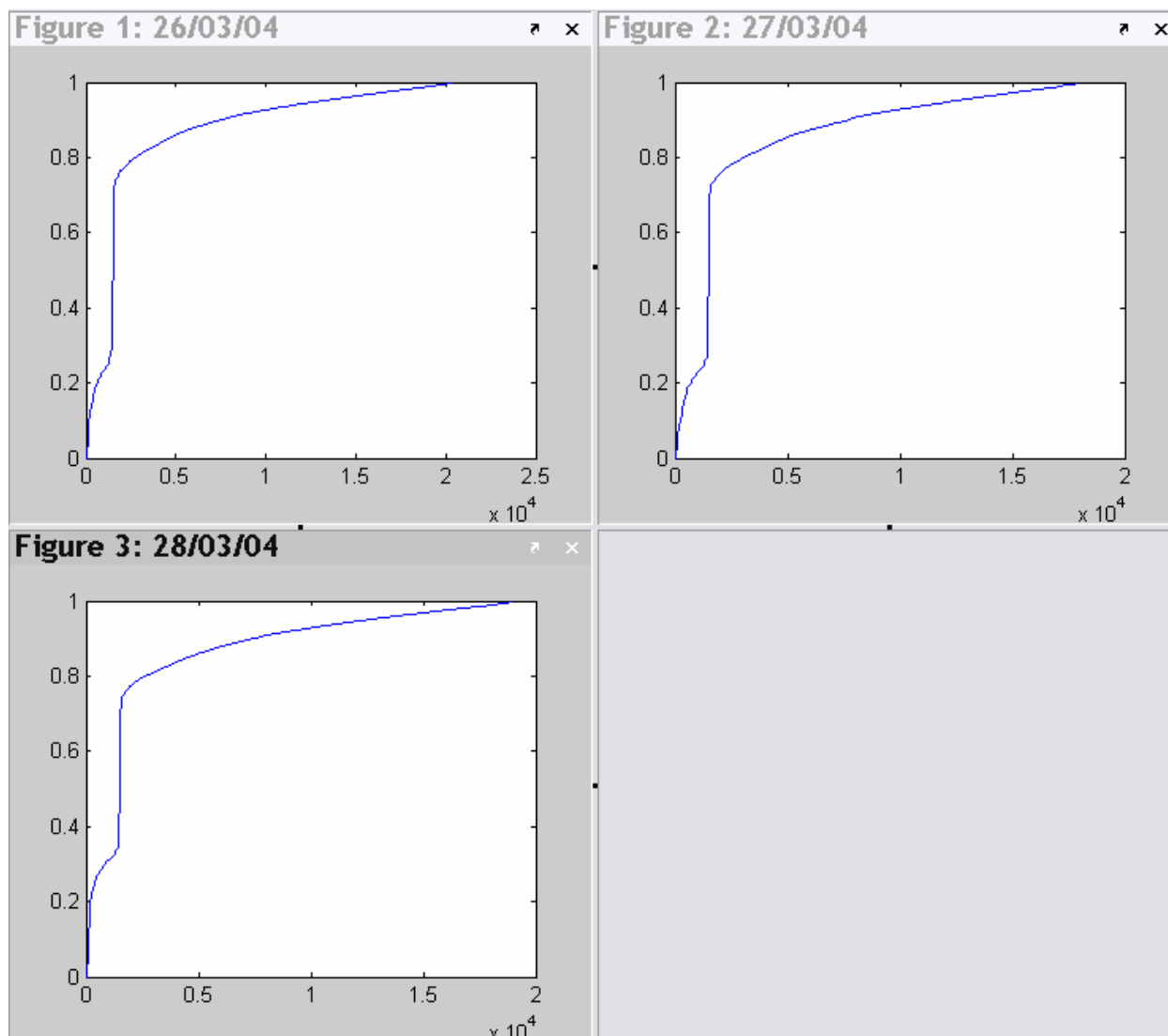


Figura 4.44. Curvas de distribución de los tamaños para el servidor PA

4.2.3. CONCLUSIONES

Las figuras 4.43 y 4.44 muestran la distribución acumulada de los tamaños de los distintos objetos transferidos por el servidor proxy RTP y PA. En todas ellas puede observarse que solo el 5% de los objetos sobrepasan los 20 KB y otro 5% no alcanzan los 100 bytes.

Existen grandes diferencias entre ambos servidores proxy para tamaños inferiores a 1500 bytes. En la tabla 3.45, elaborada a partir de algunos valores pertenecientes a las curvas de

distribución anteriormente expuestas, se observa que en el caso RTP existe una mayor concentración de objetos de pequeño tamaño. A medida que el tamaño aumenta la concertación es mayor en PA. En cualquier caso, el 80% de los objetos se encuentra entre 100 bytes y 8 KB.

	PA	RTP
100	5,64	5,93
500	17,06	48,56
1000	21,57	55,68
2000	80,51	73,22
3000	83,65	78,33
4000	87,16	82,42
5000	88,97	85,37

Tabla 4.1. Comparación de la distribución acumulada de los tamaños

4.3. FRECUENCIA DE PETICIONES

4.4.

La ley de Zipf, llamada así por el profesor de lingüística de la Universidad de Harvard, George Kingsley Zipf (1902-1950), es una curiosidad matemática que explica gran parte del comportamiento humano. El descubrimiento inicial de Zipf fue que si uno contaba el número de veces que se usaba cada palabra en distintos textos en inglés, y las ordenaba de más frecuente a menos frecuente, se cumplía que la frecuencia de la palabra i -ésima, multiplicada por i , era igual a una constante C , y la constante C dependía del texto escogido. Graficando esta curva usando una escala logarítmica en ambos ejes, ella se convierte en una recta con pendiente negativa t .

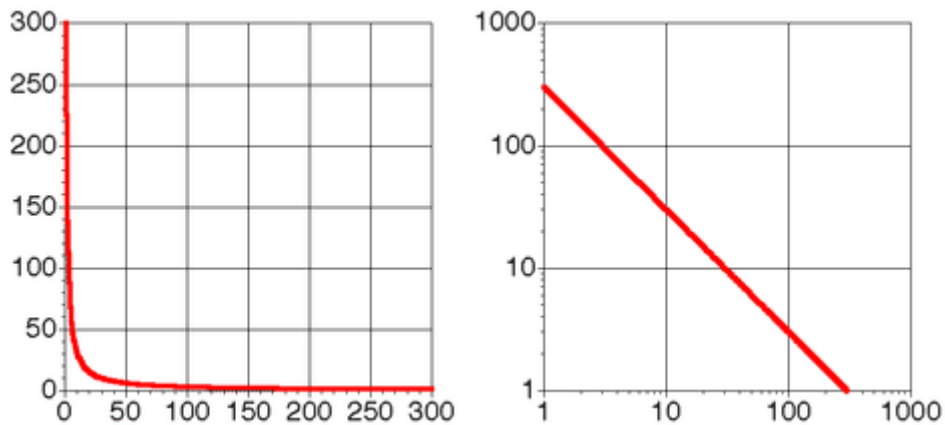


Figura 4.45. Ejemplo de gráficas de una distribución Zipf

El propósito de este capítulo es determinar si las frecuencias de peticiones a los distintos objetos siguen una distribución Zipf. Para ello se ha hecho uso de la opción ‘Gráfica de las URL más visitadas’ de la aplicación principal, expuesta en el [apartado 2.5.2.6](#). En este caso las gráficas generadas tienen escala logarítmica en los dos ejes, por lo que deberán guardar semejanza con la gráfica de la derecha de la figura 3.56 (si es que se sigue dicha distribución). En el eje X se representan las URL, y en el eje Y la frecuencia con la que se ha solicitado. Por cuestiones de requerimientos de memoria y complejidad, es imposible realizar el cálculo de frecuencias para el servidor RTP (cuyos registros de acceso superan en la mayoría de los casos los dos millones de trazas), es por esto que restringiremos este estudio únicamente al servidor PA.

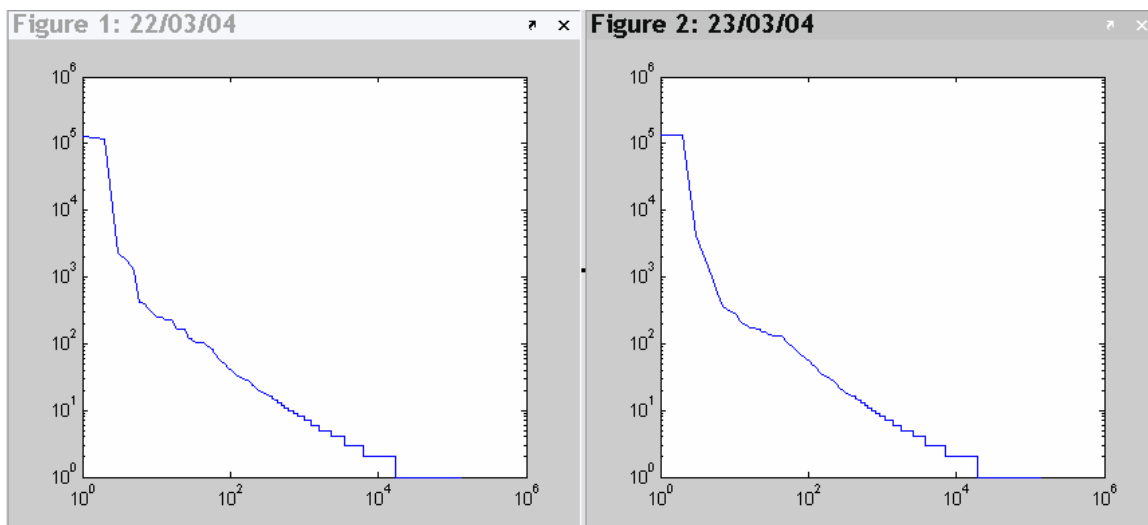


Figura 4.46. Frecuencia de peticiones para el 22/03/04 y 23/03/04

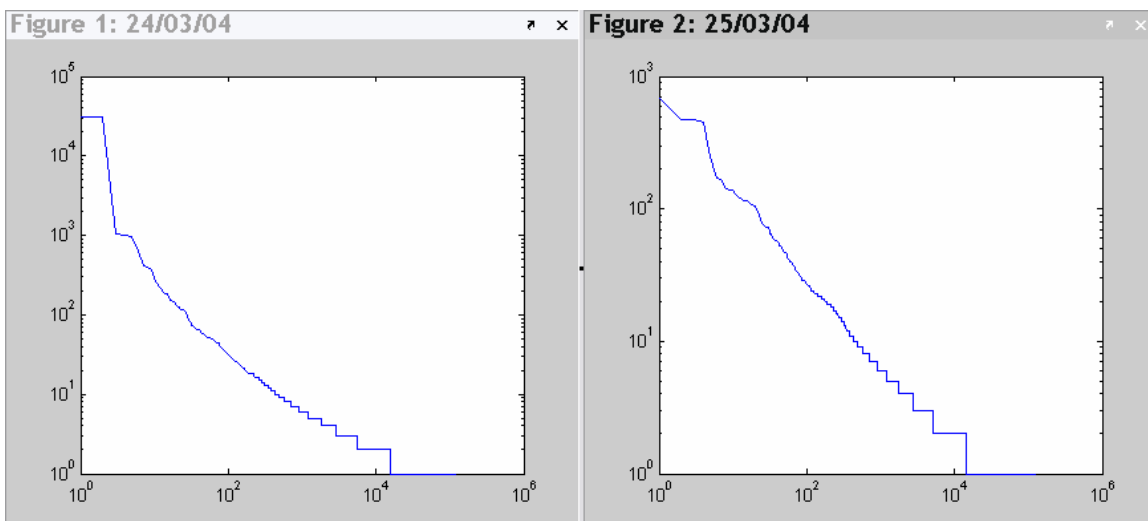


Figura 4.47. Frecuencia de peticiones para el 23/03/04 y 24/03/04

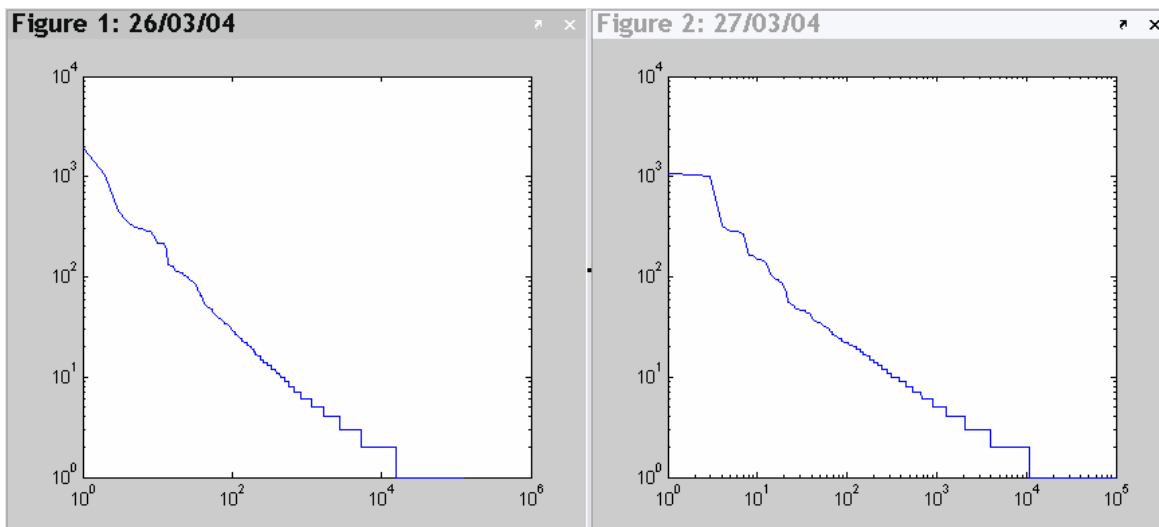


Figura 4.48. Frecuencia de peticiones para el 26/03/04 y 27/03/04

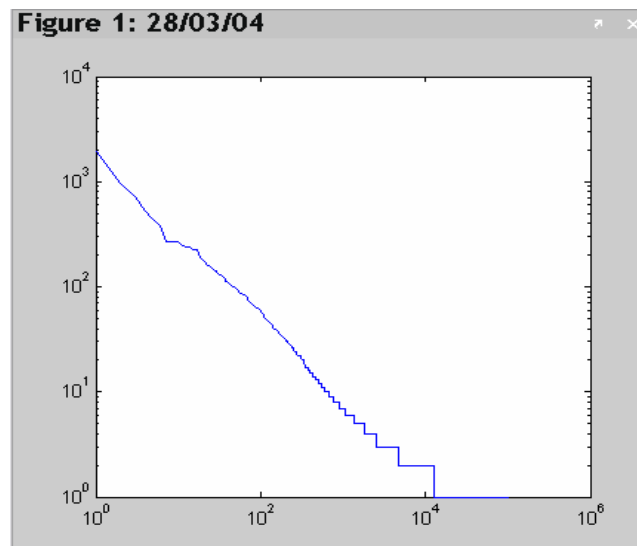


Figura 4.49. Frecuencia de peticiones para el 28/03/04

4.3.1. CONCLUSIONES

Claramente no todos los documentos Web son iguales. Algunos son extremadamente populares, a los que se accede con frecuencia y en pequeños intervalos de tiempo por muchos clientes desde muchos sitios distintos. Otros, sin embargo, raramente lo son.

Ordenando la lista de los distintos documentos en orden decreciente de acuerdo al número de veces que han sido accedidos y representándolas gráficamente, se obtienen las figuras arriba expuestas.

Las gráficas correspondientes a los días 22, 23, 24, y 27 presentan una fuerte caída a partir de los dos o tres primeros valores del eje x, esto es, los dos o tres objetos más solicitados lo son con una gran diferencia con respecto al resto (este efecto no está presente en las gráficas del 26 y el 28). Existe también una ‘forma escalonada’ para los últimos valores de x, provocado por la enorme cantidad de objetos que son solicitados dos y sobre todo una vez.

El gran número de objetos que son referenciados una sola vez puede ser un efecto provocado por la duración finita de las trazas estudiadas, o por algo tan inocente como el borrado o renombrado del documento Web. No obstante, el efecto está presente en todas y cada una de las trazas estudiadas. Este 15% de objetos que son referenciados una sola vez ocuparían las proxy cache de manera inútil. Se hace por tanto necesario técnicas para ‘purgar’ estos objetos. Aplicar *timeouts* a los objetos cacheados podría ser una buena solución. Otra aproximación podría ser la de hacer *caching* de objetos con una segunda referencia en un especificado periodo de tiempo. No obstante, esta última solución conllevaría guardar el estado de objetos que aún no han sido ni siquiera cacheados, con lo que estaríamos posiblemente incurriendo en una técnica más pesada que cachear desde la primera referencia.

En resumen, se podría concluir, que salvo el primer y último tramo de la gráfica, la ley de Zipf explica el comportamiento con el que se solicitan objetos en la Web.

CAPÍTULO 5: Simulaciones proxy-caches

El continuo crecimiento de la Web puede llevar a una sobrecarga de los servidores Web y la congestión de Internet, que tienen como resultado final el incremento en la latencia en la recepción por parte del usuario de los documentos Web.

Existe un tipo especial de servidores Web a los que se les conoce como proxy. Un proxy es capaz de actuar como servidor y como cliente. El proxy actúa como un agente, representa al servidor de cara a un cliente, y a un cliente de cara a un servidor. Un proxy acepta peticiones de un cliente y las remite a un servidor. Una vez que el proxy recibe la respuesta del servidor, éste la pasa a los clientes. Originalmente los proxy fueron diseñados para proporcionar acceso al Web a usuarios de una red privada, quienes podían sólo acceder al Web a través de un firewall. Los servidores proxy hacen más que sólo emitir respuestas del tipo HTTP. Los proxy de Web pueden ser configurados para que actúen como caches. De ser el caso se convierten en proxy-cache. En sistemas distribuidos a gran escala como la Web, el *caching* es un tema clave para obtener un buen rendimiento. La idea básica del *caching* es simple: guardar de manera local los documentos que han sido más recientemente (o frecuentemente) extraídos para ser usados en futuras peticiones, evitando así tener que extraerlos desde los sitios de origen. De esta manera el *caching* mejora la velocidad de acceso y disminuye el tráfico en la red. También contribuye en la reducción de la carga en los servidores originales e incrementa la disponibilidad de la información debido a que la replica.

La efectividad del *caching* puede ser medida principalmente por dos cantidades: La tasa de aciertos (o hits), la cual define el número de peticiones satisfechas por la cache, y el tamaño de los documentos, que es importante tener como métrica debido a la alta variabilidad que presentan, a esto llamamos la tasa de aciertos en bytes (o el byte hit ratio).

Debido a la alta complejidad y consumo de recursos del proceso de simulación, restringiremos el este estudio al servidor PA.

El estudio se ha centrado en las clases y tipos de objetos que más aportan al tráfico, bien por un alto porcentaje en la frecuencia de peticiones, bien por un alto porcentaje de bytes que

aportan al tamaño total. Estos son imágenes y paginas estáticas, para el caso de las clases, y los gif para el caso de los tipos.

5.1. POLÍTICAS DE REEMPLAZO

Las políticas de reemplazo a estudiar en este apartado serán las expuestas a continuación.

5.1.1. POLÍTICA DE REEMPLAZO LRU (LEAST RECENTLY USED)

Esta política de reemplazo se basa en lo siguiente:

1. Probablemente los objetos más recientemente usados seguirán usándose en el futuro cercano.
2. Probablemente los objetos no muy recientemente usados no se usarán mucho en el futuro cercano.

Con esta política se desaloja de la caché el objeto que tiene más tiempo sin usarse. En la implementación software se hace uso del archivo de timestamps (out0) generado por la aplicación de pre-procesamiento para almacenar el momento en que se solicita un objeto. Si es necesario un reemplazo, se sustituye entonces aquel objeto que tenga la referencia más antigua.

5.1.2. POLÍTICA DE REEMPLAZO LFU (LEAST FREQUENTLY USED)

Esta política de reemplazo se basa en lo siguiente:

1. Probablemente los objetos más usados seguirán usándose en el futuro.
2. Cuando varios objetos han sido usados con la misma frecuencia y son candidatos al reemplazo se sustituye aquel objeto que tenga la referencia más antigua.

En la implementación software cada objeto posee un contador que se incrementa cada vez que el objeto ha sido referenciado. Se sustituye aquel que tenga su contador más bajo. En caso de existir varios candidatos al reemplazo, se procede igual que en el caso LRU.

5.1.3. POLÍTICA DE REEMPLAZO HÍBRIDA

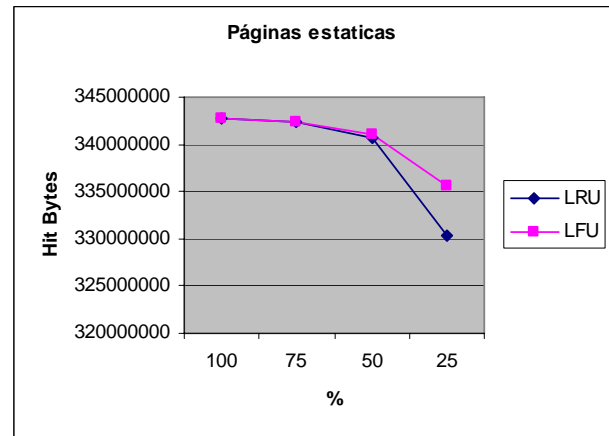
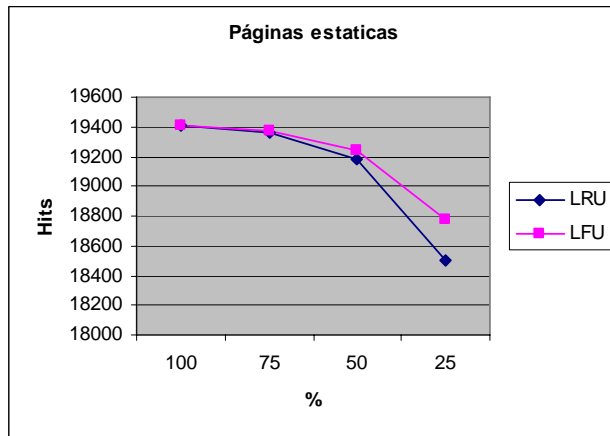
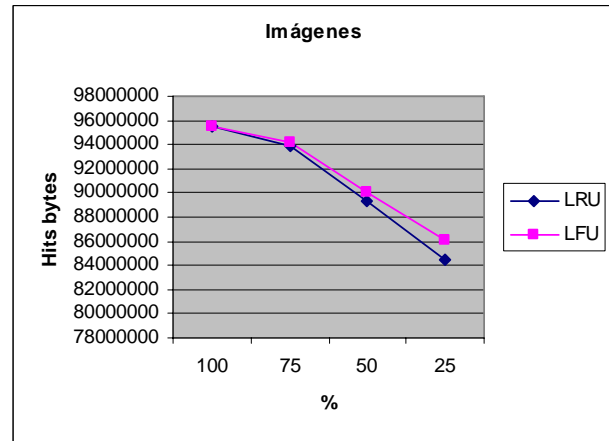
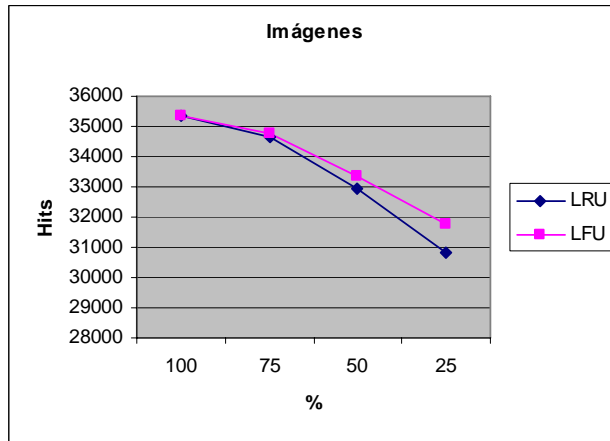
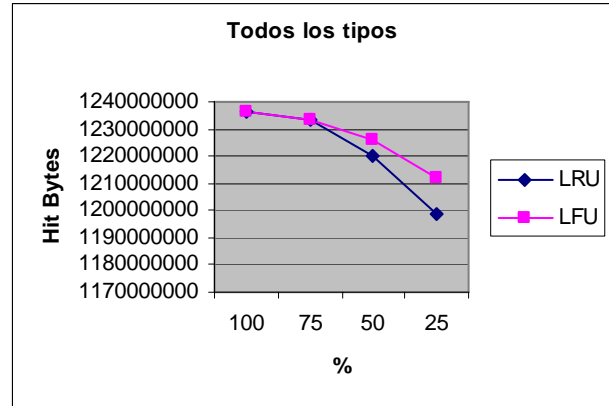
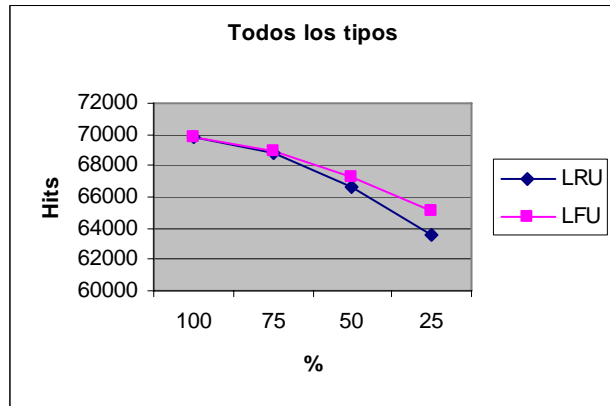
Consiste en aplicar una u otra de las políticas anteriores dependiendo de la clase de objeto a reemplazar. Para este estudio hemos considerado las clases que más aportan al tráfico: imágenes y documentos HTML. De esta forma tendremos que representar las siguientes combinaciones de proxy caches:

1. con política de reemplazo LFU para los documentos HTML y LRU para imágenes
2. con política de reemplazo LRU para los documentos HTML y LFU para imágenes.

No se ha requerido de ninguna implementación software adicional para simular dicha política, sino que simplemente se han simulado por separado cada política de reemplazo con cada clase de objeto y posteriormente se han sumado sus hit-rates y sus hit-byte-rates.

5.2. RESULTADOS

En este apartado se muestran las gráficas de rendimiento de las proxy cache para 100%, el 75%, el 50%, y 25% del tamaño mínimo necesario para que la cache tuviera el 100% de la efectividad. Este rendimiento es expresado en el número de hits (gráficas de la derecha) y de hits byte rate (gráficas de la izquierda). Este estudio, que comprende los datos de toda la semana para el servidor PA, se ha realizado para las imágenes, para las páginas estáticas, para los objetos de tipo gif, y para todos los objetos en general sin distinción de tipo ni clases. Además se muestran las gráficas de rendimiento para el estudio de simulación híbrida, que como anteriormente se explicaba, divide una misma cache en dos, aplicando distintas políticas de reemplazo dependiendo de la clase de objeto que va a almacenar.



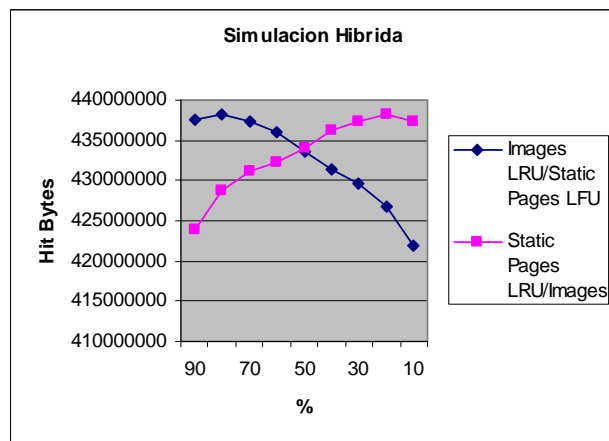
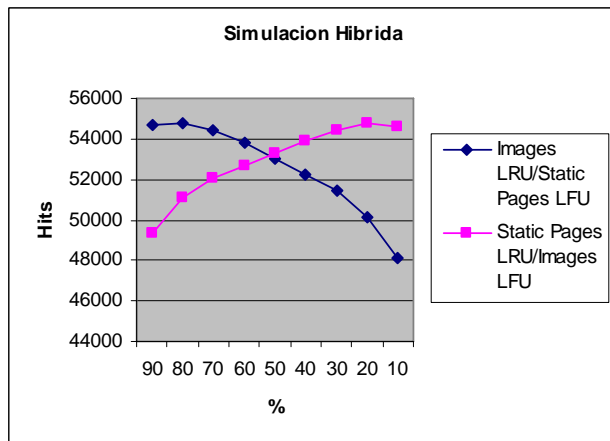
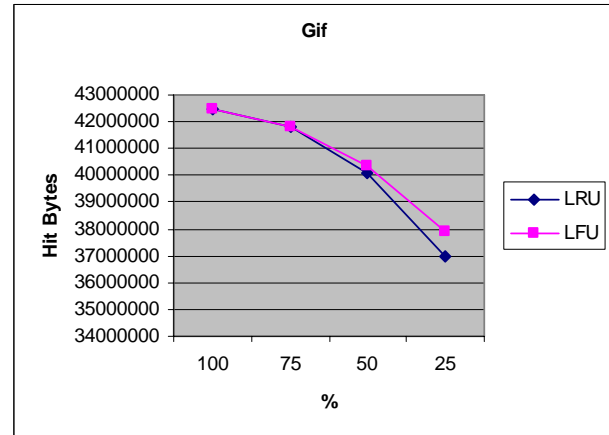
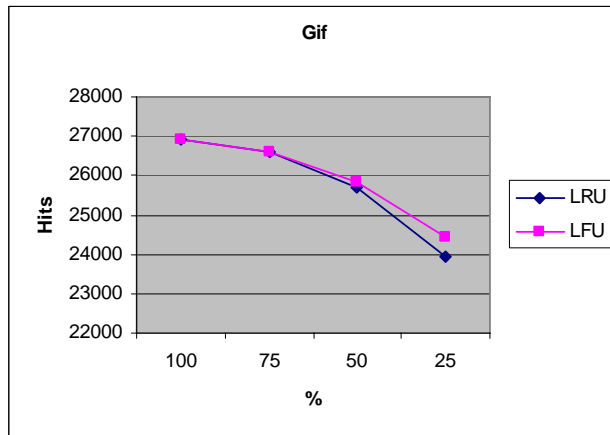
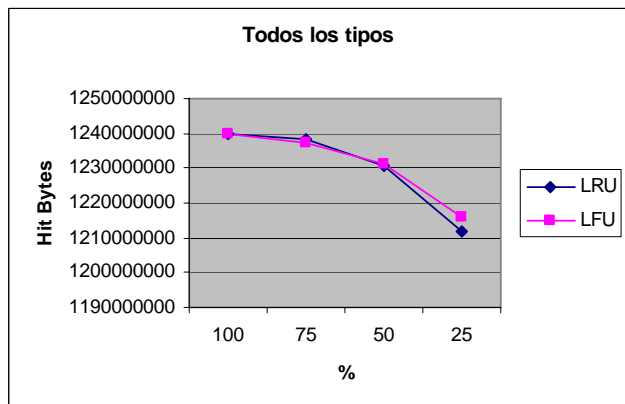
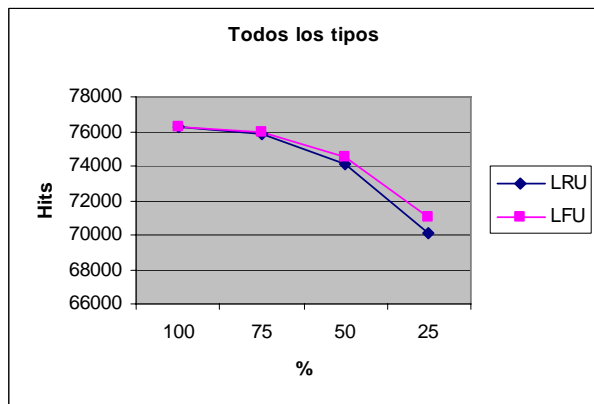
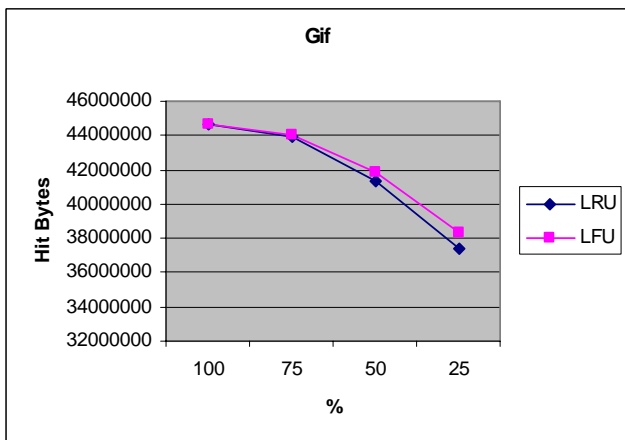
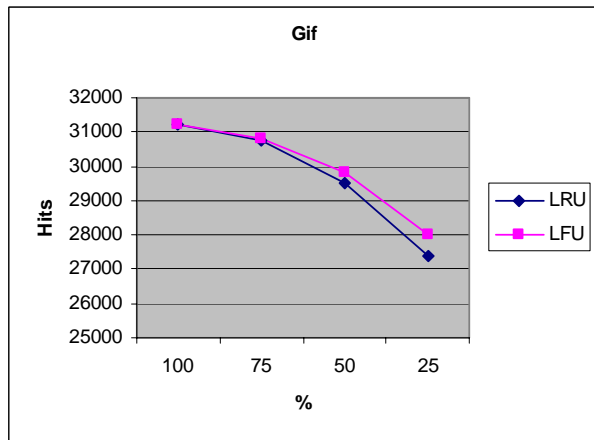
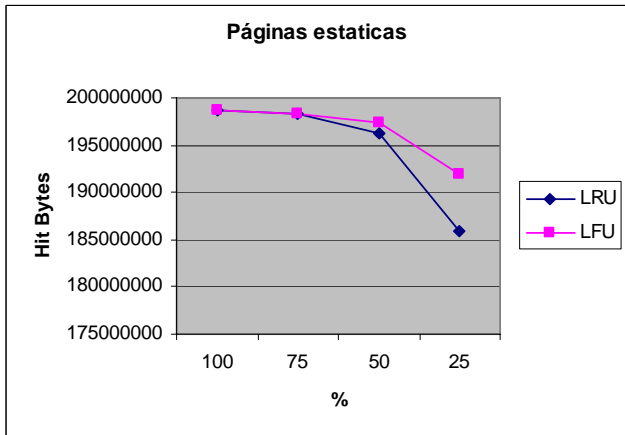
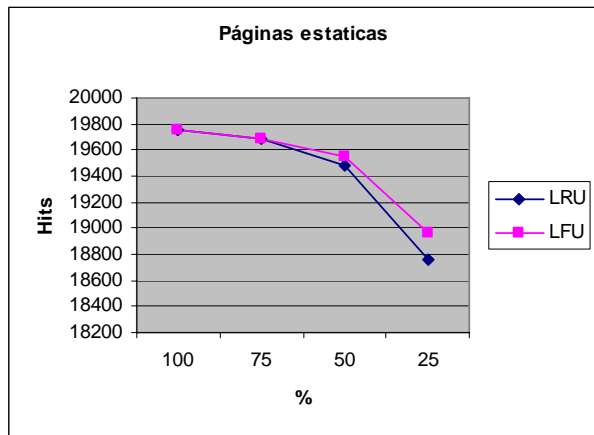
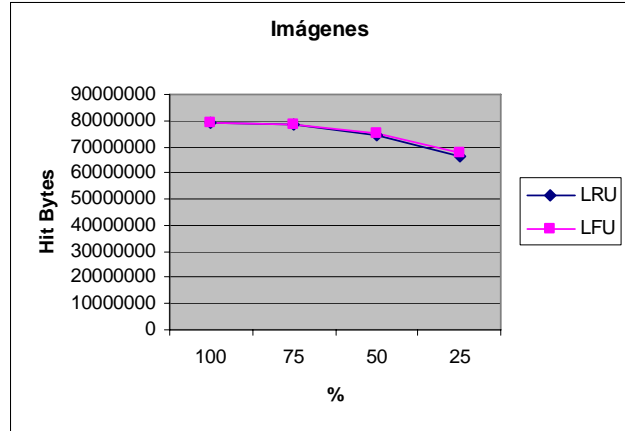
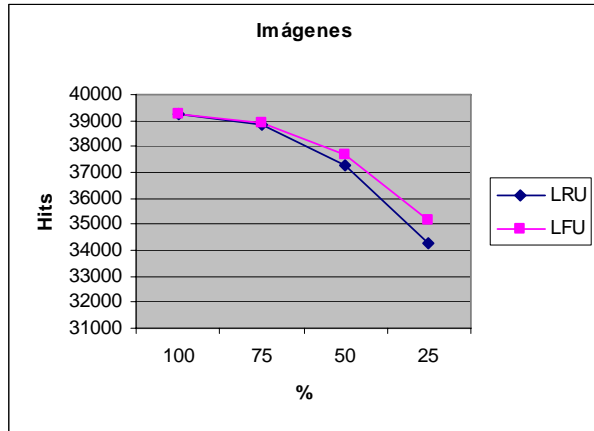


Figura 5.1. Rendimiento de la proxy-cache para 22/03/04





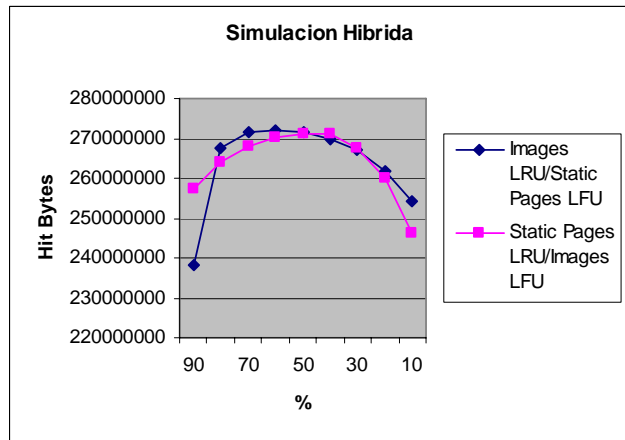
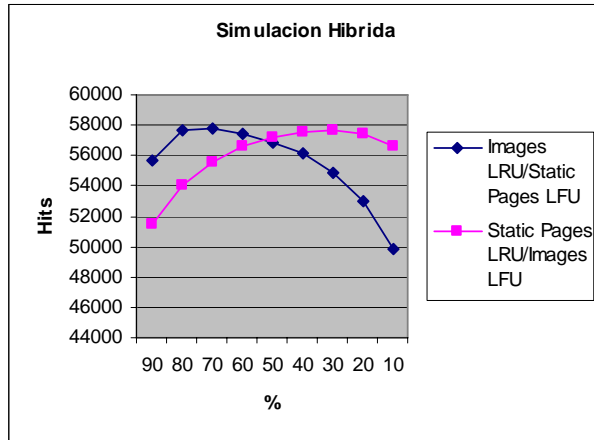
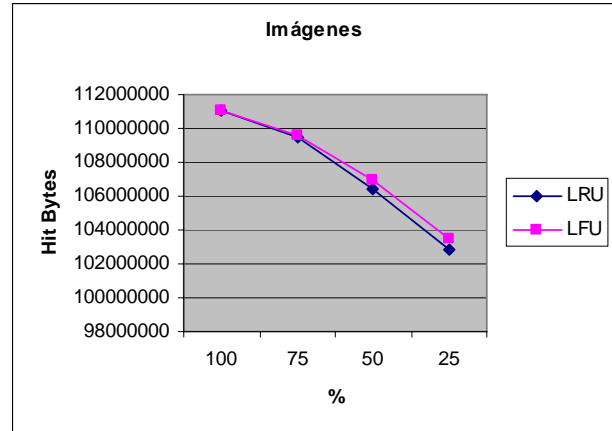
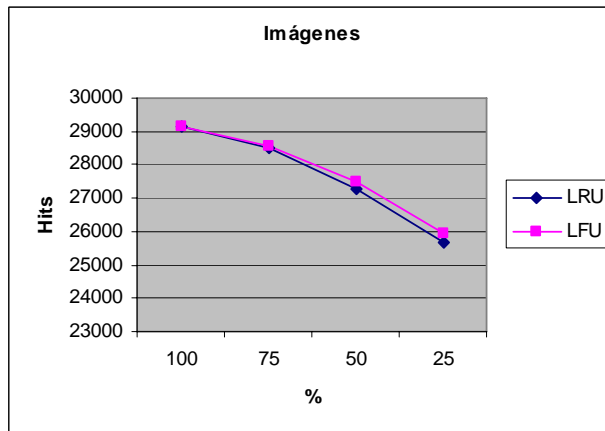
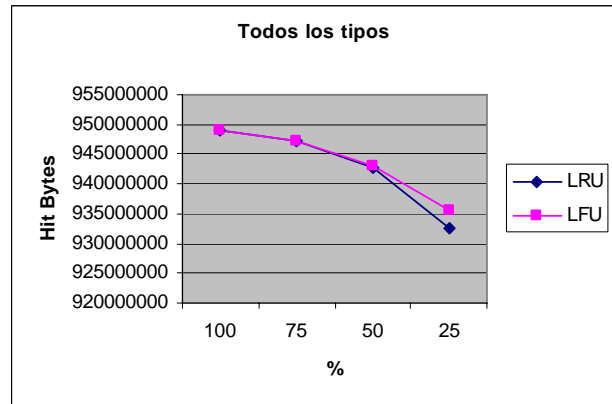
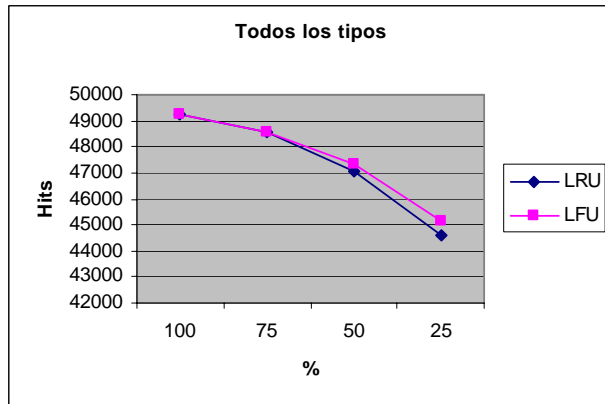


Figura 5.2. Rendimiento de la proxy-cache para 23/03/04



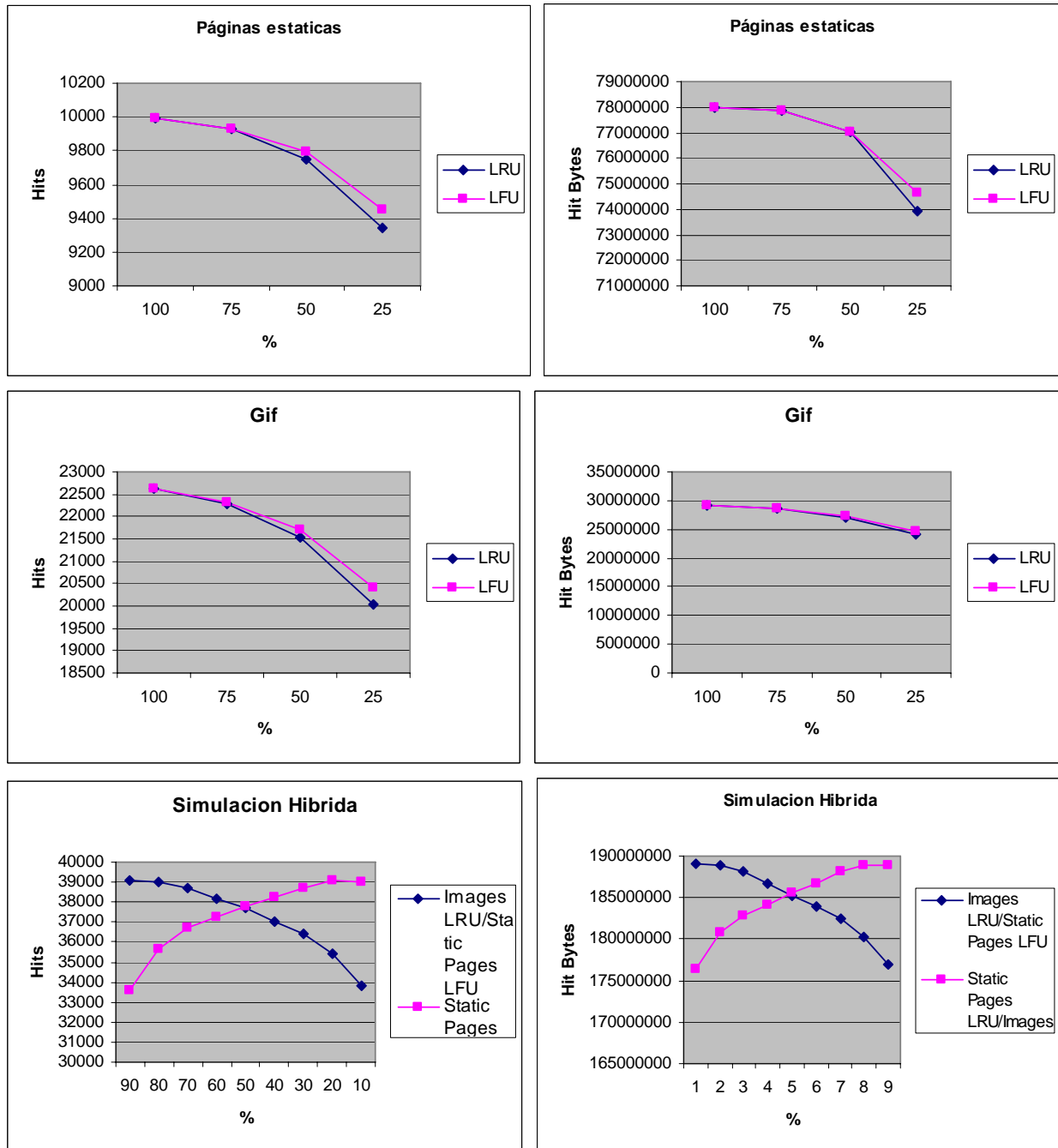
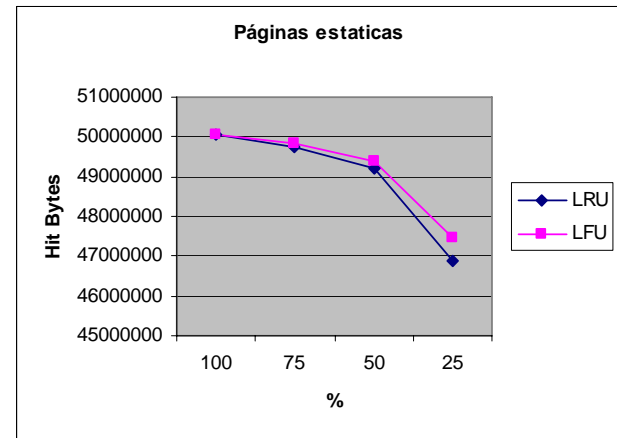
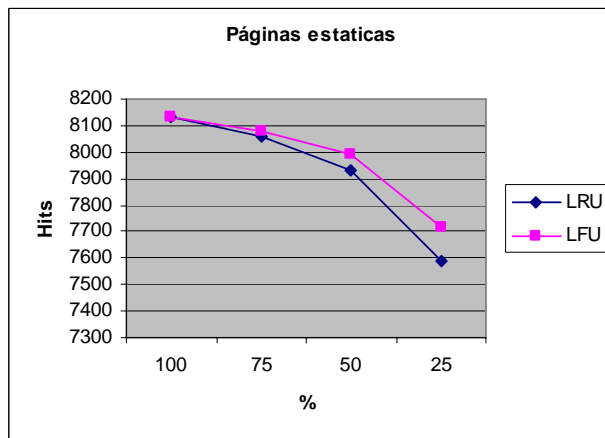
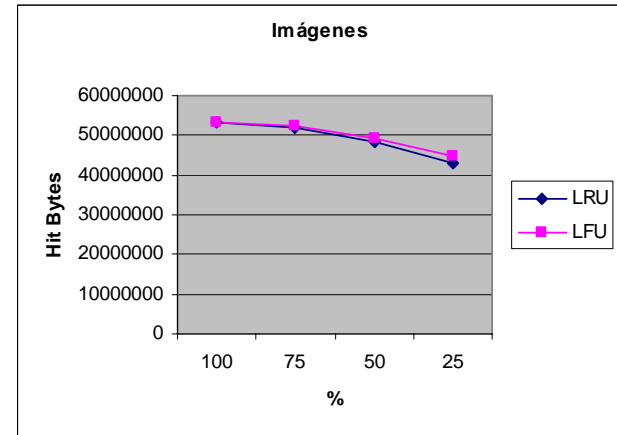
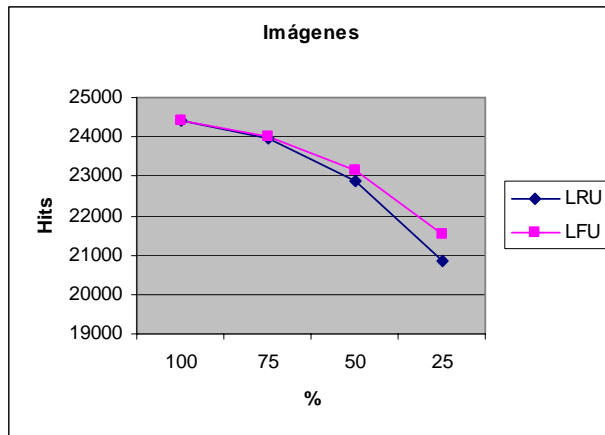
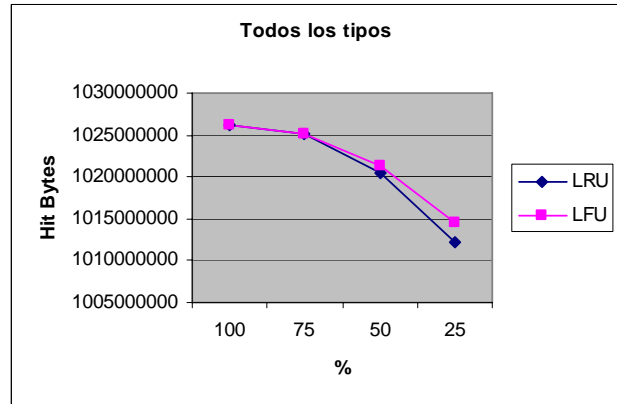
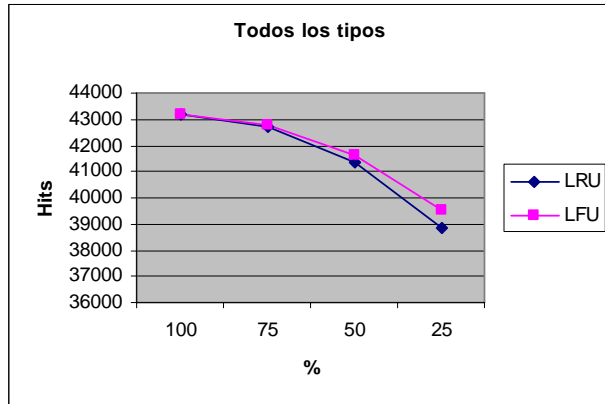


Figura 5.3. Rendimiento de la proxy-cache para 24/03/04



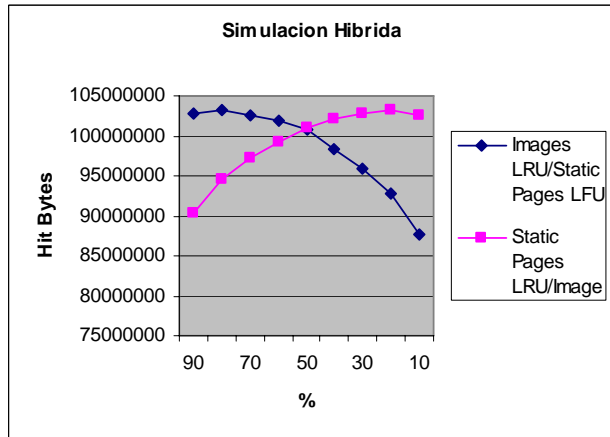
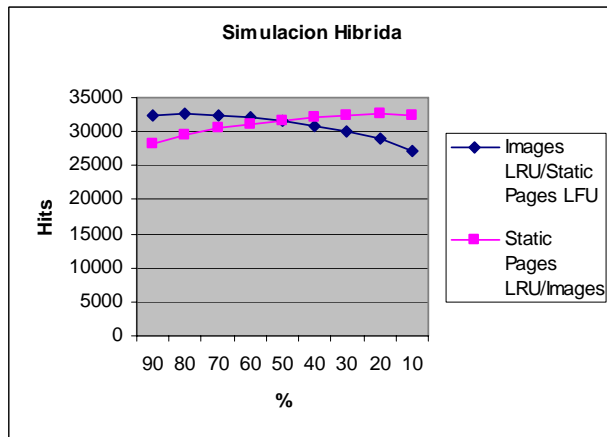
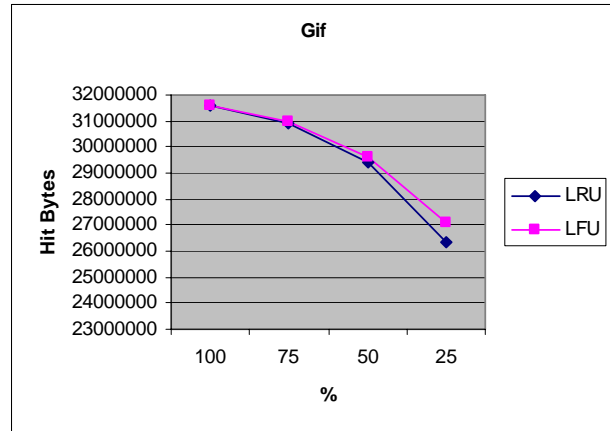
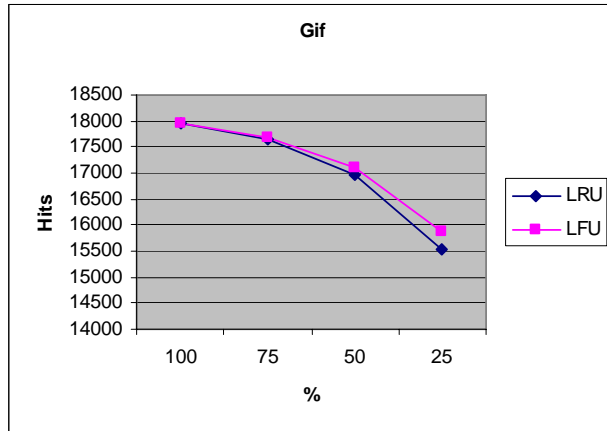
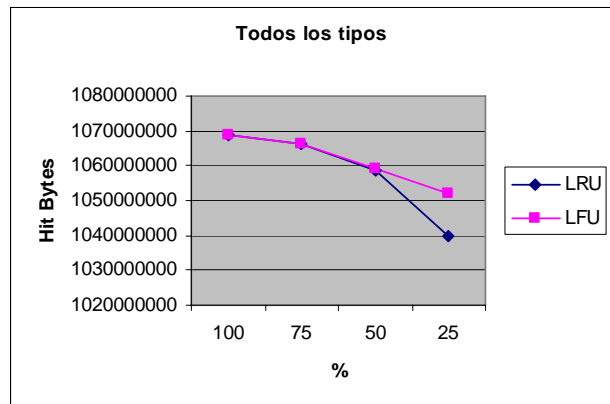
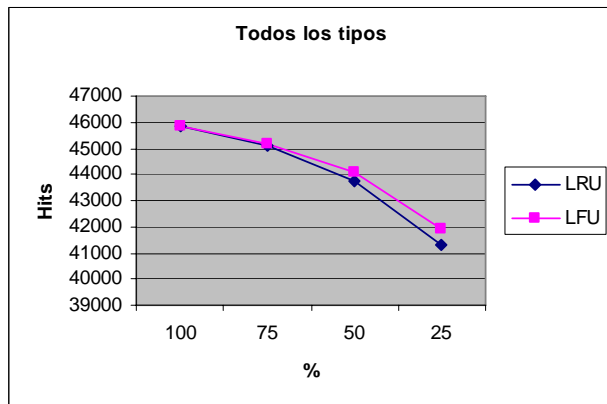
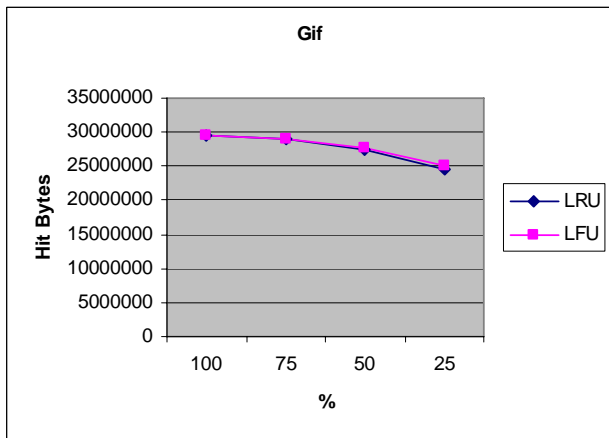
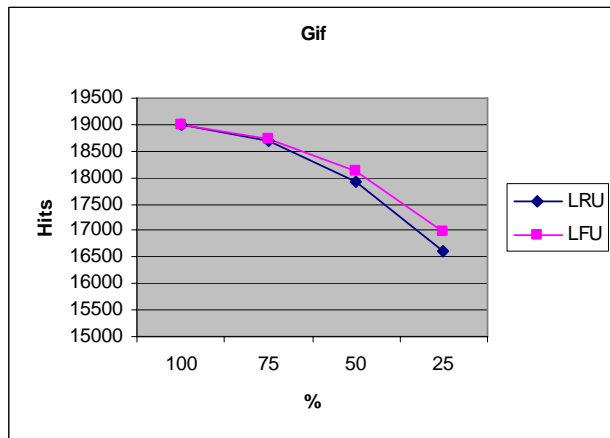
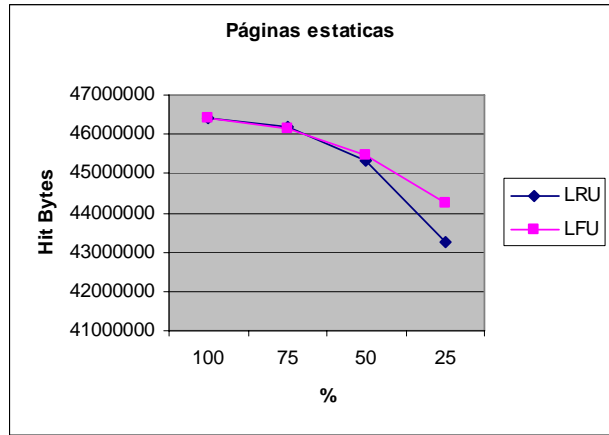
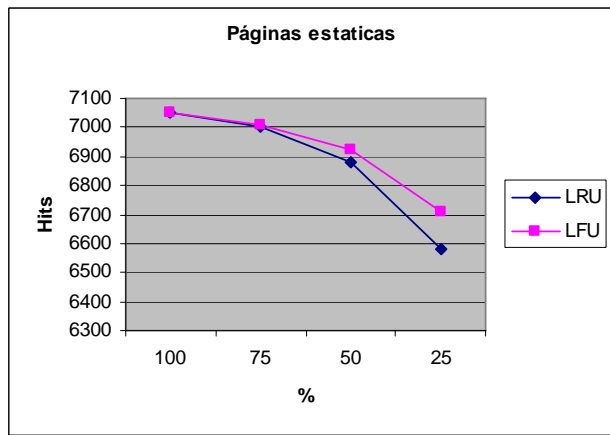
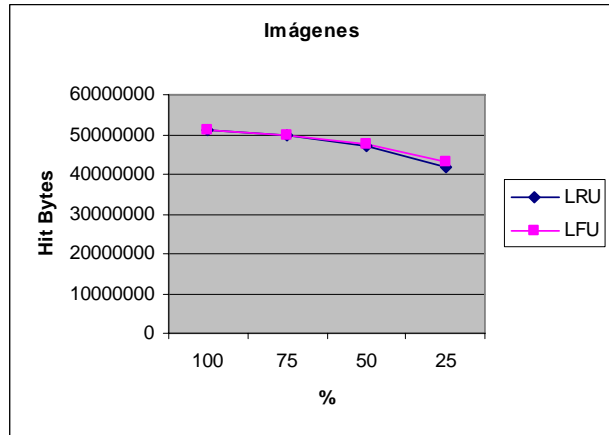
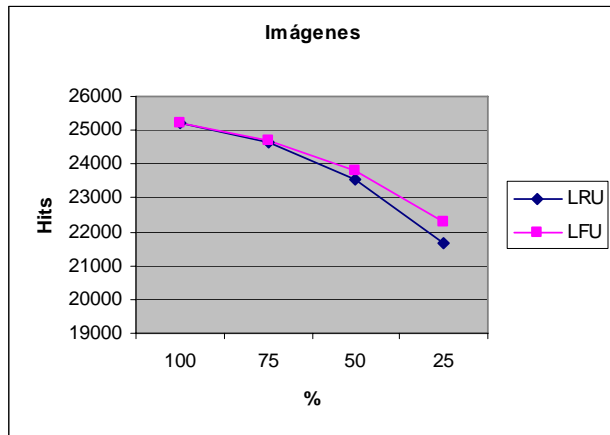


Figura 5.4. Rendimiento de la proxy-cache para 25/03/04





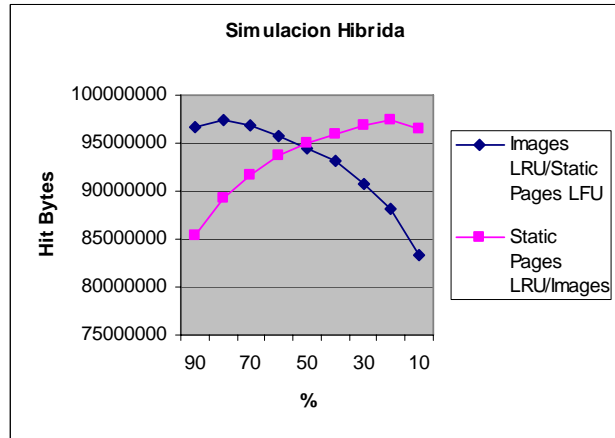
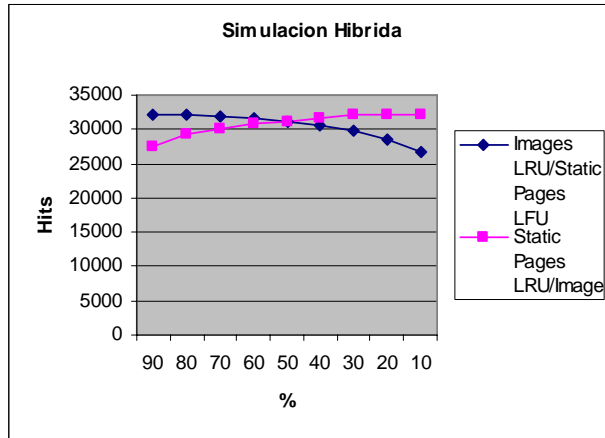
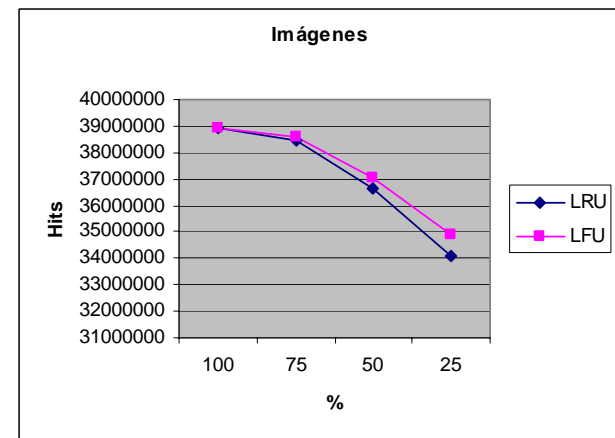
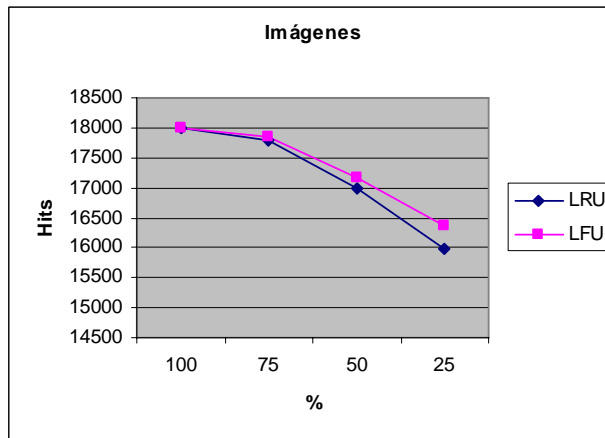
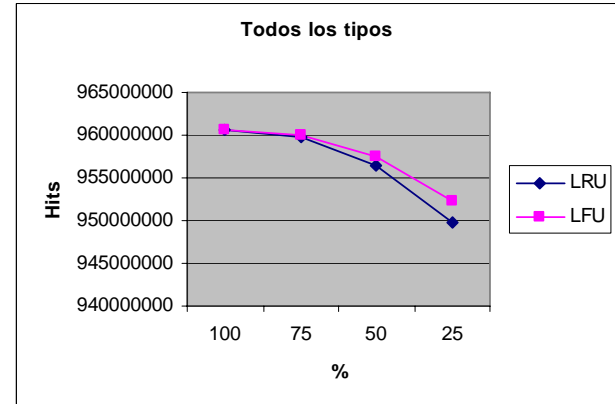
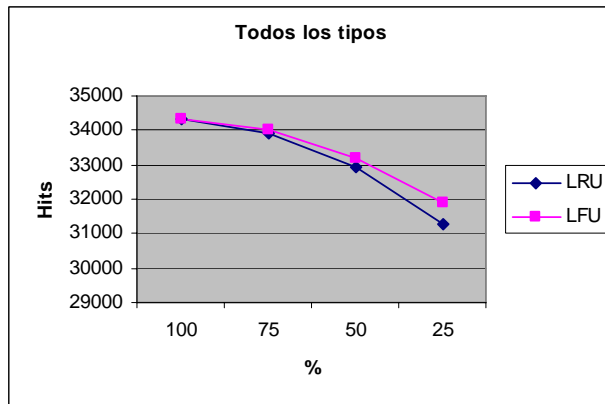


Figura 5.5. Rendimiento de la proxy-cache para 26/03/04



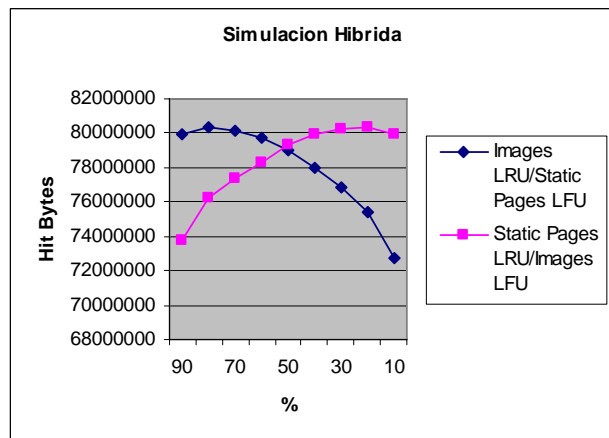
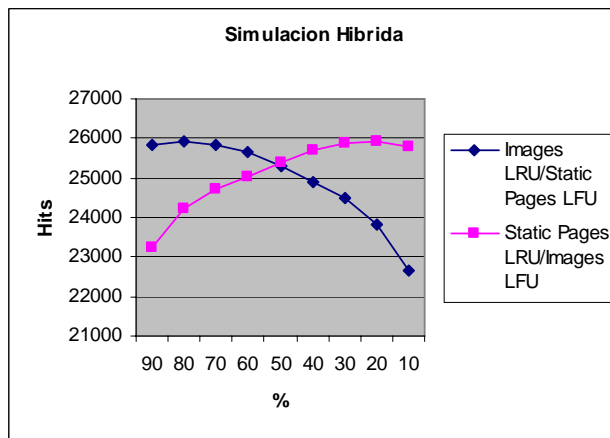
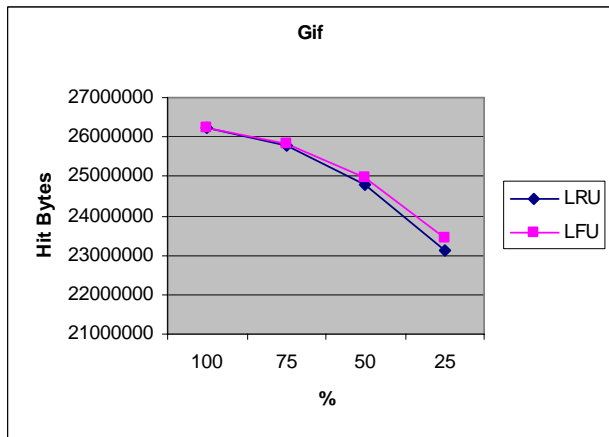
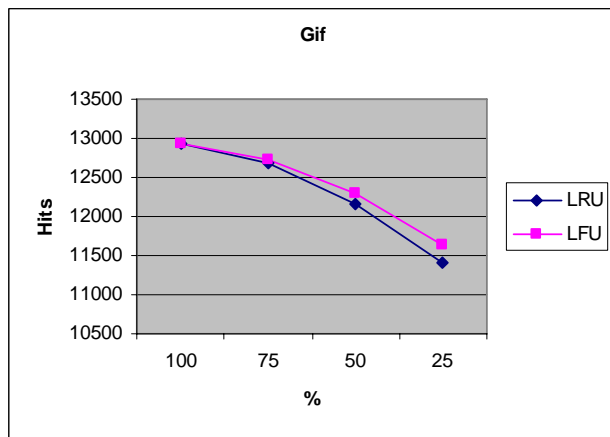
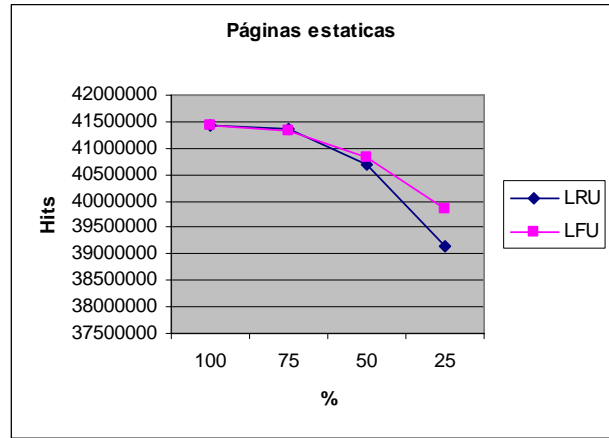
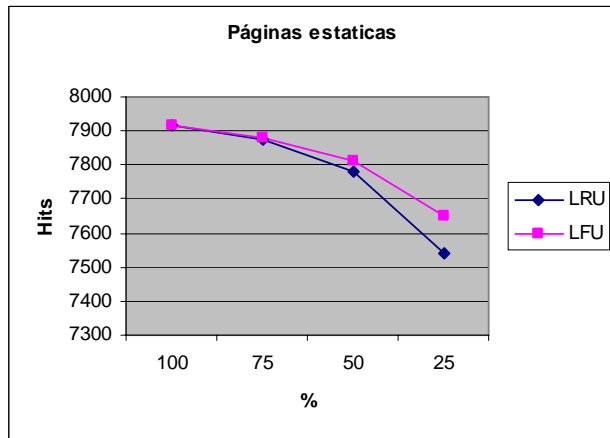
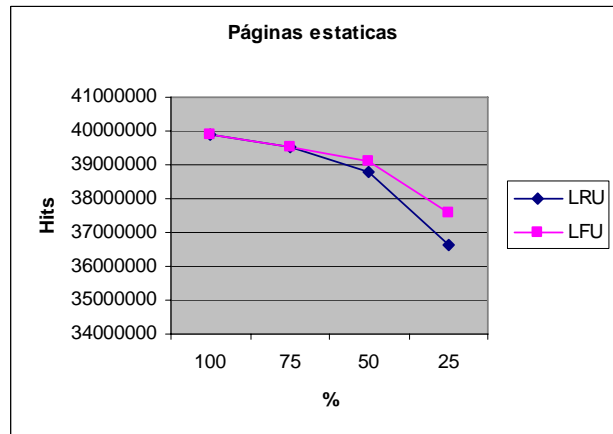
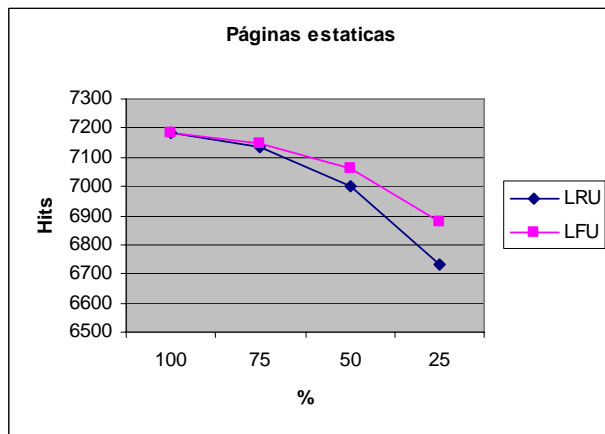
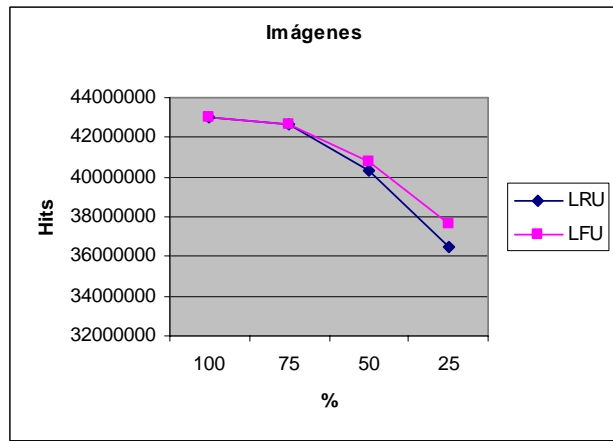
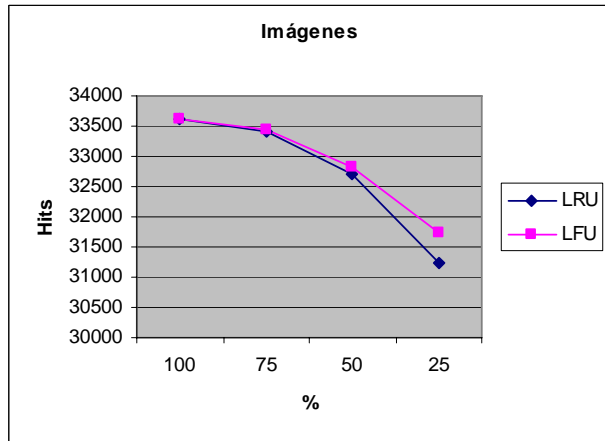
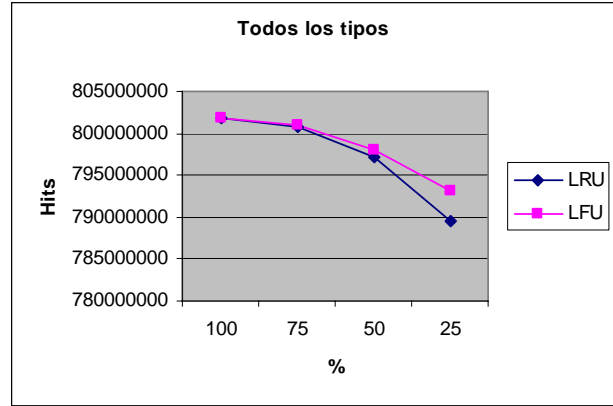
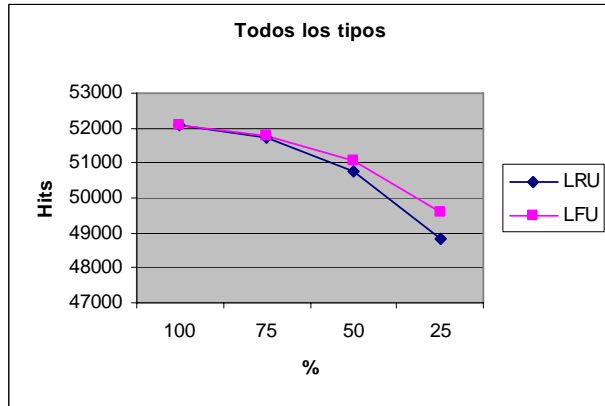


Figura 5.6. Rendimiento de la proxy-cache para 27/03/04



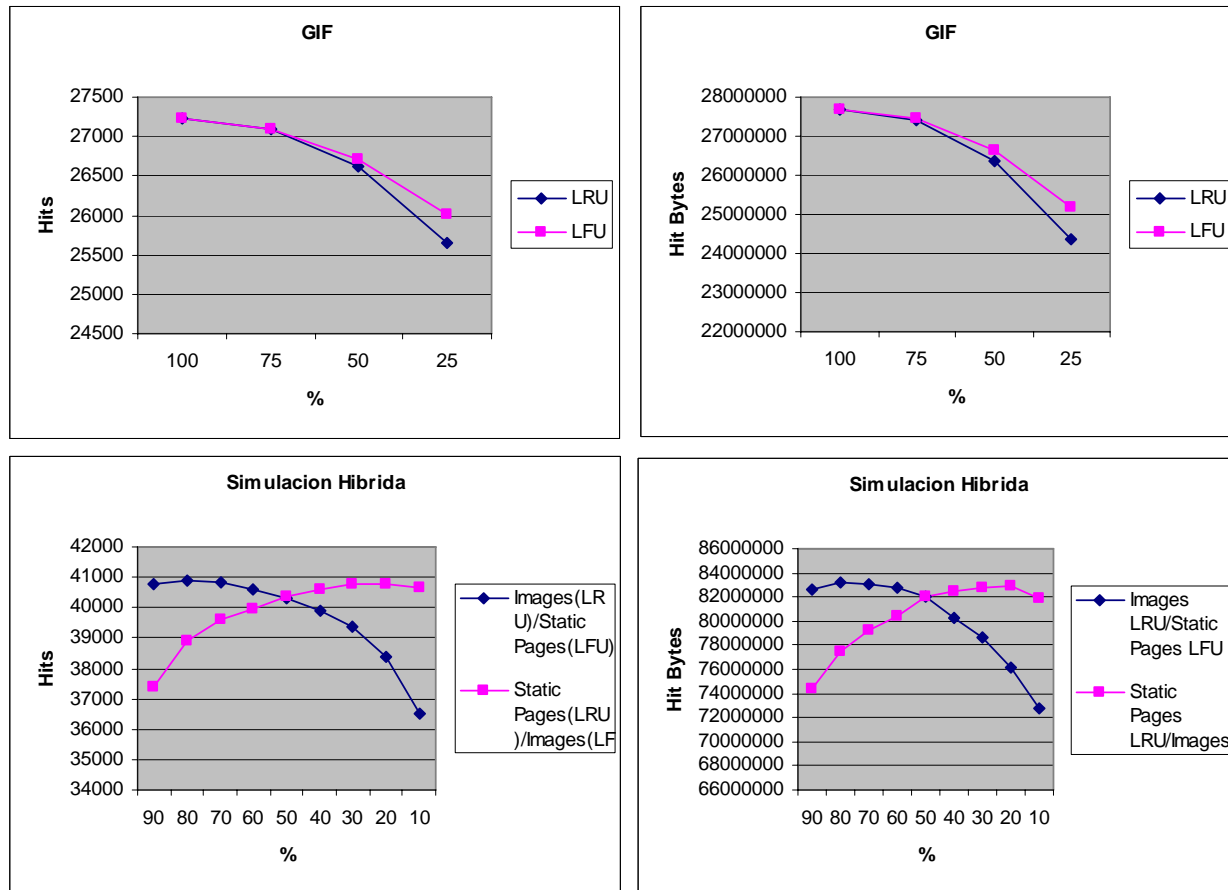


Figura 5.7. Rendimiento de la proxy-cache para 28/03/04

5.3. CONCLUSIONES

Todas las gráficas de rendimiento anteriores muestran la misma forma. De igual manera, en el caso de las imágenes, páginas estáticas, gif en particular, o sin diferenciar tipos o clases de objetos, las gráficas de rendimiento desvelan una mayor eficiencia en el caso de las proxy cache con política de reemplazo LFU, lo cual se traduce en la escasa presencia de la localidad temporal en las peticiones. Esta diferencia en el rendimiento se hace mayor cuanto menores son los tamaños de los proxy cache. Así encontramos que los proxy cache con reemplazo LFU, tienen una eficacia superior en muchos casos al 2% con respecto a la política LRU cuando el tamaño de estas es del 25% del total.

El estudio también desvela pequeñas diferencias en el rendimiento de ambas políticas dependiendo de las clases o tipos de objetos que se cachean. Así, en la tabla 5.1 se muestran las

medias de los porcentajes de caída de eficiencia de la política LRU con respecto a la LFU, en función de la clase o tipo de objeto cacheado. Se observa que para el caso de las imágenes estas diferencias de eficiencia se hacen más evidentes.

Todos Tipos	1,62
Imágenes	2,28
html	1,57
gif	1,97

Tabla 5.1. Diferencias de eficiencia entre proxy caches LFU y LRU

Por otro lado, en el caso de la simulación híbrida la mejor solución es aquella en la que las páginas estáticas son tratadas con una política de reemplazo LRU y las imágenes con la LFU. El caso óptimo resulta cuando el 20% del tamaño se dedica al almacenamiento de las páginas estáticas y el resto para las imágenes. El resultado no sorprende si se tiene en cuenta que las imágenes aportan mayor número de bytes a las trazas que las páginas estáticas (tal y como se mostraba en el [apartado 3.2.1.3](#)), y que el almacenamiento para este tipo de objeto sea la política LFU, ya que esta resulta más eficiente, tal y como se comenta algunas líneas más arriba.

CAPÍTULO 6: Conclusiones y líneas futuras

Los principales invariantes arrojados en el estudio estadístico de las trazas, realizado en el [capítulo 3](#), son los siguientes:

1. El método GET constituye más del 99% de las peticiones.
2. Las imágenes y las páginas estáticas forman el 40% del tamaño total, y entorno al 75% de las peticiones,
3. y en concreto los objetos que más aportan al tráfico son los jpeg y los html.
4. Los grupos con mayor tasa de transferencia son los ‘otros’ y los sonidos, mientras que en el otro extremo estarían los ‘http’ y las páginas dinámicas.
5. A pesar del hecho de que los navegadores Web proveen soporte para el uso de objetos multimedia, tales como sonido y video, estos a penas aportan el 8% al total del tráfico.

No obstante, la continúa proliferación de páginas cada vez más complejas, que incluso se hacen más fáciles con herramientas tales como CGI scripts y Java, podrían cambiar considerablemente los resultados estadísticos obtenidos.

En el [apartado 4.1](#) se estudio el espacio entre peticiones a un mismo objeto, llamando la atención la intensa pero breve popularidad de los objetos estudiados. Era por ejemplo el caso de los dos objetos más populares, los cuales eran referenciados una media de 10500 veces al día, popularidad que solo duro hasta el tercer día del estudio, a partir del cual no volvió a ser referenciado; o el caso del cuarto objeto que mantuvo su popularidad solo el primer día.

Por otro lado, el estudio de las curvas de distribución de los tamaños de los objetos realizado en el [apartado 4.2](#), arrojaba que eran muy pocos los de pequeño tamaño (< 100 bytes), la gran mayoría de estos oscilaban entre 100 bytes y 8 KB, y pocos ($< 5\%$) eran mayores de 20 KB. Partiendo de estos resultados, podría resultar interesante contrastar los resultado obtenidos en este trabajo con los procedentes de modificar el software de simulación para añadir a la cache un limite en el tamaño del objeto. Así, se evitaría cachear objetos de gran tamaño, poco utilizados y que ocuparan gran cantidad de espacio; o demasiado pequeños, ya que estos quitan pocos bytes a transferir desde el servidor.

Los resultados obtenidos en el apartado del estudio de las frecuencias de peticiones, realizados en el [apartado 4.3](#), podrían también ser aplicados al capítulo de simulación de proxy-cache. El estudio concluía que más de un 15% de los objetos se solicitaban una sola vez, y por tanto, el cacheado de estos resultaría inútil. Aplicar *timeouts* a los objetos cacheados podría ser una buena solución. Otra aproximación podría ser la de hacer *caching* de objetos con una segunda referencia en un especificado periodo de tiempo. No obstante, esta última solución conllevaría guardar el estado de objetos que aún no han sido ni siquiera cacheados, con lo que estaríamos posiblemente incurriendo en una técnica más pesada que cachear desde la primera referencia.

En el estudio del rendimiento de las proxy caches, desarrollado en el [capítulo 5](#), se destacaba la LFU como la política de reemplazo más eficiente, proporcionando casi un 3% de eficiencia más que para el caso LRU, concluyendo que el principio de localidad temporal temporal. También se mostró que estas diferencias variaban dependiendo del tipo y clase de objeto cacheado. Así, para el caso de las imágenes esta diferencia se hacía máxima.

Algunos estudios llevados a cabo en este proyecto han quedado limitados al uso de los datos obtenidos a partir del servidor proxy PA. El gran volumen de datos aportados por RTP ha hecho que las herramientas ‘distUrls2’ y ‘frecUrls’ tuvieran unos tiempos de ejecución y niveles de consumo de recursos intratables.

En consecuencia, el estudio de la aproximación de las frecuencias de peticiones de objetos en la Web a una distribución Zipf y el estudio de las distancias entre peticiones han sido solo realizados para el caso del servidor proxy más pequeño (PA).

Rediseñar los algoritmos de búsqueda y ordenación utilizados en estas herramientas y la utilización de memoria secundaria para cálculos intermedios, podrían ser soluciones alternativas a la solución más obvia de aumentar los recursos hardware.

Por otro lado, utilizar otras características de medición e implementar más políticas de reemplazo son algunas de las posibles ampliaciones para el estudio del rendimiento de las proxy cache, visto en el capítulo 5.

APÉNDICE A: CLASIFICACIÓN DE EXTENSIONES

Esta clasificación de extensiones ha sido la implementada en la aplicación de pre-procesamiento y troceado de trazas, explicada en el [apartado 2.1](#).

Imágenes	Jpg 1	jpeg 2	Jpe 3	Gif 4	Tif 5	Bmp 6	Pcx 7	Png 8
	Ico 9	Psp 10	Tiff 11	Vst 12	Pct 13	Pcd 14	Xbm 15	Ras 16
	Pgm 17	Ppm 18	Pic 19					

Compri- midos	Zip 37	Gz 38	Rar 39	Ace 40	Tar 41	Z 42	Tgz 43	Gbd2zip 44
	Bz2 45	Zoo 46						

Aplica- ciones	Exe 73	Cab 74	Class 75	Ini 76	Dll 77	Bin 78	Jar 79	War 80
	Ear 81	Bat 82	Cmd 83					

Vídeos	Mpg 109	Mpeg 110	Mov 111	Avi 112	Ram 113	Qt 114	Mp2 115	Mpa 116
	Wmv 117	Rm 118	Vcs 119	Asx 120	Dcr 121	Swf 122		

Sonídos	Wav 145	Mp3 146	Mid 147	Midi 148	Fpl 149	Au 150	Snd 151	Asf 152
	Rmi 153	Basic 154	Wma 155	Rpm 156	Stm 157	Sw 158	Abs 159	Aif 160
	Aiff 161	Aob 162	Axr 163	Cda 164	Dsm 165	Gig 166		

	Htm	Html	Xml	Xhtml	Shtml	Dhtml	Wml	Sgml
--	-----	------	-----	-------	-------	-------	-----	------

HTML	181	182	183	184	185	186	187	188
	Xstl 189	Map 190						

HTML dinámico	Php 217	Asp 218	Cgi 219	Jhtml 220	Ashx 221	Php3 222	Shtm 223	shtml 224
	phtml 225	Php4 226	Fhtml 227	Jsp 228	Aspx 229			

Texto	Txt 253	Cfg 254	Plain 255
--------------	------------	------------	--------------

HTTP	http 289
-------------	-------------

Otros	Cdf 325	Rdf 326	Main 327	Iso 328	Grp 329	Dat 330	Dt 331	Gsz 332
	Flg 333	X32 334	X86 335	Trg 336	Mds 337	Tag 338	Cb 339	Ads 340
	Cdx 341	Usr 342						

Scripts	Js 361	Vbs 362	Vb 363
----------------	-----------	------------	-----------

Docu- mentos	Xls 397	Pps 398	Srf 399	Cfm 400	Adp 401	Doc 402	Ppt 403	Slm 404
	Sld 405	Wp 406	Wpd 407	Wpg 408	Ans 409	Bfx 410	Cdk 411	Chi 412
	Dox 413	Fmp 414	Gno 415	Hed 416	Pdf 417	Ps 418	Rtf 419	Src 420

ircache	Irc 433
----------------	------------

Presenta- ción	Xsl 469	Css 470
---------------------------	------------	------------

Descono- cidos	-2	-1
---------------------------	----	----

BIBLIOGRAFÍA

- Martin F. Arlitt, “Web Server Workload Characterization: The Search for Invariants”, University of Saskatchewan, 1996
- Ghaleb Abdulla, “Analysis and Modeling of World Wide Web Traffic”, Thesis doctoral, Faculty of the Virginia Polytechnic Institute, 1998
- T. Berners-Lee, “Hypertext Transfer Protocol 1.0”, RFC 1945, 1996
- T. Berners-Lee, “Hypertext Transfer Protocol 1.1”, RFC 2616, 1999
- John O. Limb, “A Behavioural Model of Web Traffic”, Georgia institute of Technology, 2000
- Kevin W. Froese, “The Effect of Client Caching on File Server Workloads”, University of Saskatchewan, 1999