FEUP
Universidade do Porto
Faculdade de Engenharia

Mestrado em
Redes e Serviços de
Comunicação

Relatório do trabalho prático

Análise de Tráfego de Internet

Análise e Modelização de Sistemas e Redes

Luís Filipe Pinto de Almeida Teixeira
Pedro Miguel Machado Carvalho

27 de Fevereiro de 2003

Índice

Resposta 1

Na Tabela 1 estão representados os principais protocolos identificados no traço. O Gráfico 1 é uma representação gráfica dos valores da tabela.

Protocolo	Nº de pacotes	Percentagem de pacotes	Nº de bytes	Percentagem de bytes
TCP	114328	76.409%	59709407	67.407%
UDP	34219	22.87%	28762772	32.471%
ICMP	975	0.652%	90988	0.103%
Outros	105	0.069%	17592	0.019%

Tabela 1 - Distribuição dos pacotes/bytes por protocolo

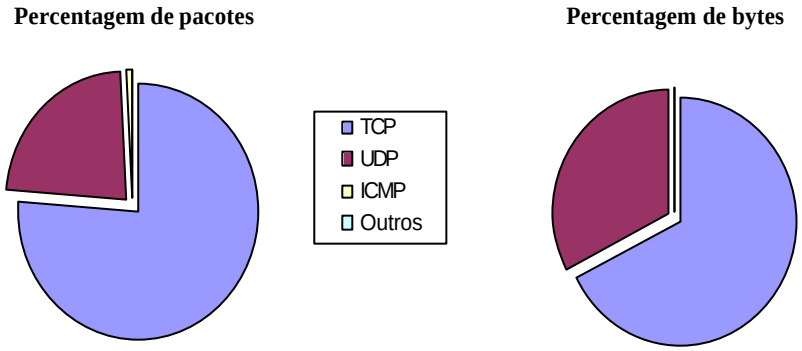

Gráfico 1 - Distribuição dos pacotes/bytes por protocolo

Os protocolos mais significativos são o TCP, UDP e ICMP, pertencendo a maior percentagem ao TCP. Também no que diz respeito ao número de bytes, o TCP possui a maior percentagem, no entanto, as diferenças são menores. Esta diferença é justificada pelo comprimento médio dos pacotes de cada protocolo, indicados na Tabela 2. Pode observar-se que o comprimento médio dos pacotes TCP é bastante inferior (aproximadamente 48%) ao comprimento médio dos pacotes UDP. Uma vez que o TCP necessita de estabelecer e terminar conexões (*connection-oriented*), é necessária a troca de um grande número de pequeno pacotes para o efeito. A necessidade de confirmação de dados em TCP (protocolo fiável) também aumenta o número de pacotes de pequena dimensão. Por outro lado, o UDP (*connectionless* e não fiável) consegue enviar o mesmo

número de dados num menor número de pacotes. Estes factos justificam os resultados obtidos.

Protocolo	Comprimento médio (bytes)
TCP	522,3
UDP	840,6
ICMP	93,3

Tabela 2 - Comprimento médio dos pacotes

Resposta 2

Nos Gráficos 1 e 2 é representada a função distribuição do comprimento dos pacotes. O comprimento mínimo dos pacotes no traço é 28 bytes e o comprimento máximo é 4470 bytes.

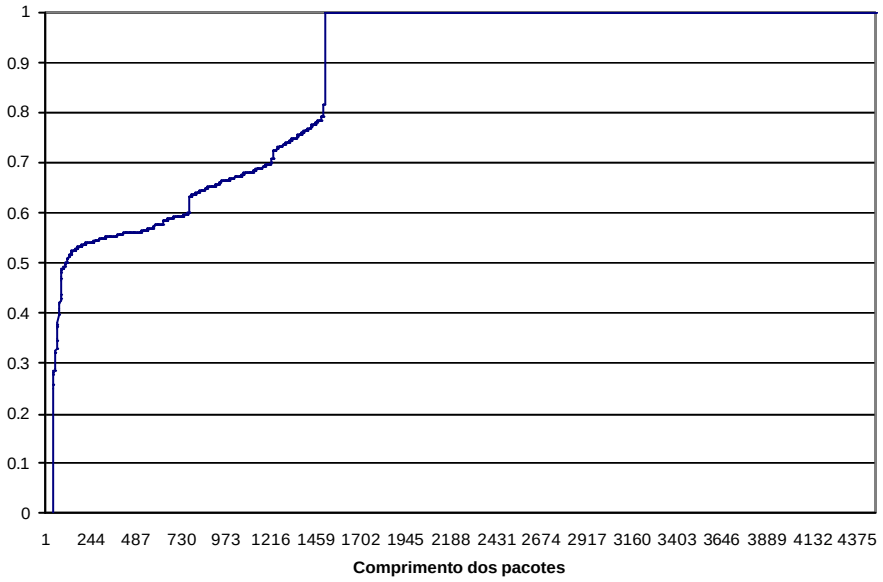

Gráfico 2 - Função distribuição do comprimento dos pacotes

Cerca de 25% dos pacotes possuem comprimentos pequenos (40 bytes) o que corresponde a pacotes TCP de estabelecimento e terminação de conexões e confirmações. Outro comprimento de pacotes com grande número de ocorrências é 1500 (cerca de 20%)

o que corresponde ao comprimento máximo de um segmento TCP (limitado pelo Maximum Transfer Unit (MTU) para a Ethernet).

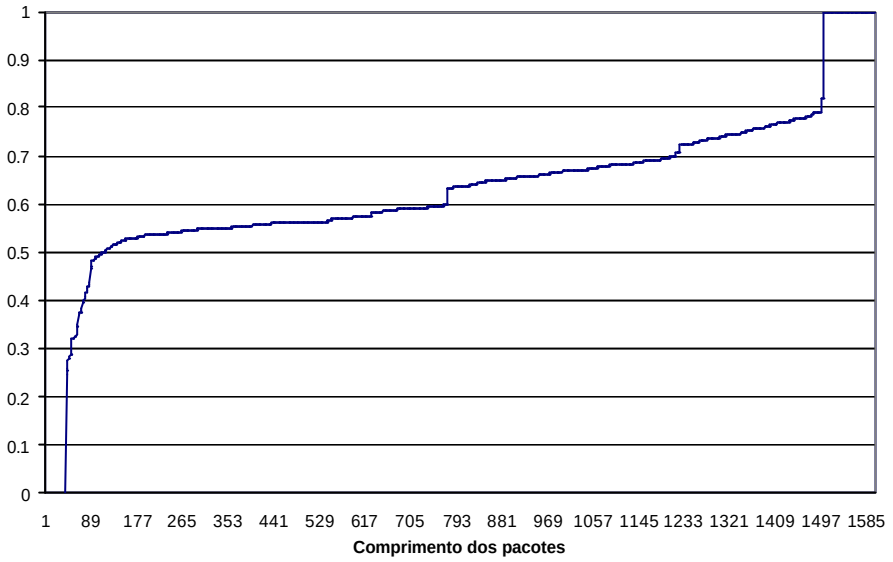

Gráfico 3 - Ampliação do gráfico anterior (Gráfico 2)

Resposta 3

Serviço	Número de pacotes	Percentagem
HTTP	28673	19.16%
FTP	17602	11.76%
NNTP	2170	1.45%
SMTP	1488	6.80%
DNS	785	0.52%
NTP	545	0.36%
Outros	10179	6.80%
Sem "porta-bem-conhecida"	88185	58.94%

Tabela 3 - Percentagem de pacotes por serviço

Na Tabela 3 encontra-se representado o número de pacotes do traço por serviço e a respectiva percentagem.

Dos serviços que utilizam uma porta "bem-conhecida", os mais utilizados são o HTTP e o FTP, que representam em conjunto 78%. Do valor total dos pacotes, os trocados entre

portas “não-conhecidas” representam cerca de 59%. Este valor indica que grande parte dos pacotes trocados na Internet são referentes a aplicações que não utilizam portas “bem-conhecidas”, e.g. FTP em modo passivo e aplicações P2P (“peer-to-peer”).

Consideram-se pacotes puramente ACKs (sem dados) aqueles que têm *flag* a indicar que é efectuada uma confirmação (flag=0x10) e que têm comprimento 40 bytes (apenas o cabeçalho). Nestas condições foram detectados **33613 pacotes puramente ACK**.

Resposta 4

Na Tabela 4 é indicado o número de fluxos no traço e o serviço associado.

Tipo de fluxo	Número de fluxos
TCP	2680
UDP	96
Outros	122
Total	**2898**

Tabela 4 - Número de fluxos por tipo

Mais uma vez o protocolo TCP é o mais significativo no traço, com 92,5% do total de fluxos. Apenas 3,3% dos fluxos são UDP. Isto deve-se ao facto de a grande maioria das aplicações utilizar o TCP como protocolo de transporte.

Resposta 5

Nos gráficos 4 e 5 são representadas as funções distribuição do tráfego. O Gráfico 5 é uma ampliação do gráfico 4 e representa os fluxos que correspondem a 90% do tráfego total.

A partir da análise dos dois gráficos anteriores pode concluir-se que uma pequena percentagem dos fluxos, representa grande parte do tráfego total. Neste traço, um *router* necessitaria de armazenar informação sobre **102 fluxos** para manter um registo de **90% do tráfego**.

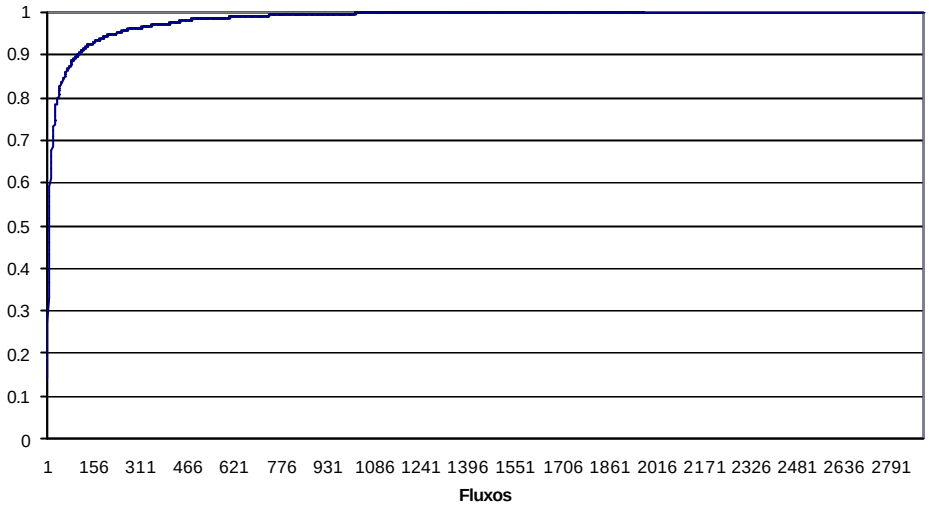

Gráfico 4 – Função distribuição do tráfego gerado pelos fluxos

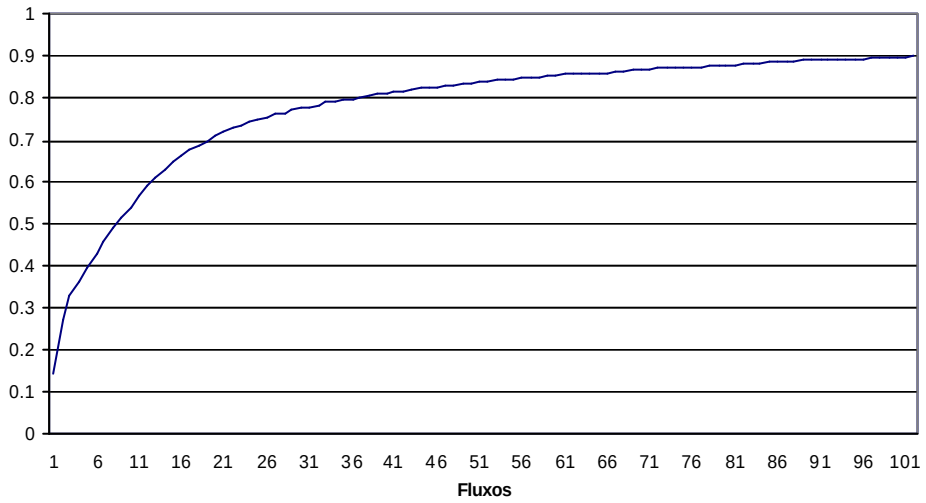

Gráfico 5 - Ampliação do gráfico anterior (Gráfico 4)

Resposta 6

Optou-se por não se fazer a representação desta resposta sobre o gráfico da resposta 5, uma vez que utilizam grandezas diferentes para as coordenadas. Na resposta anterior é representada a percentagem de tráfego em função do fluxo enquanto que nestes gráficos é feita a representação do débito em função do tempo. Os Gráficos 6, 7 e 8 representam o débito com intervalos de contagem de 10ms, 100ms e 1s, respectivamente.

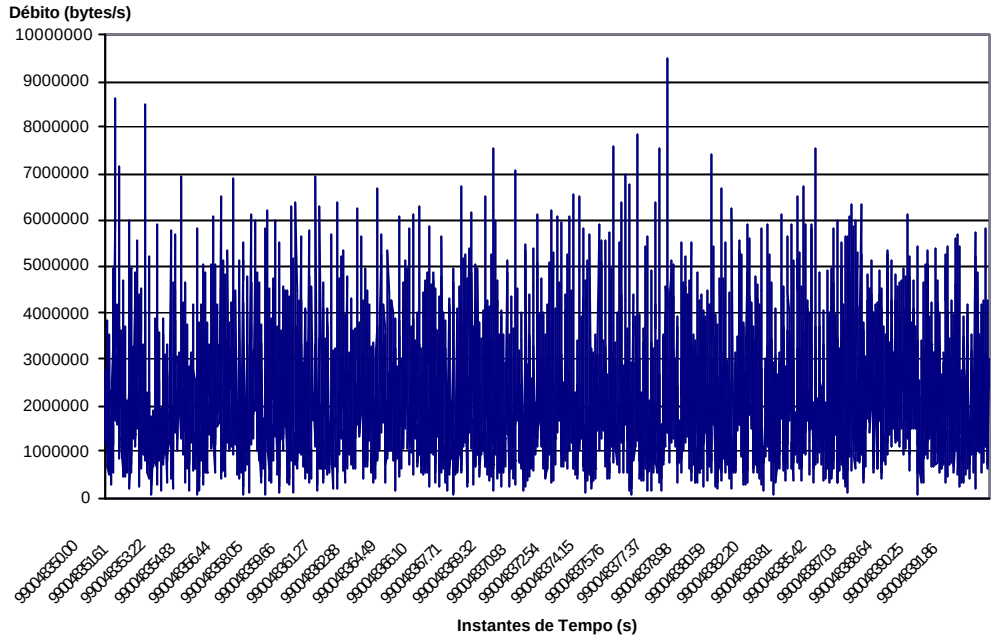

Gráfico 6 - Débito do traço com 10ms de intervalo de contagem

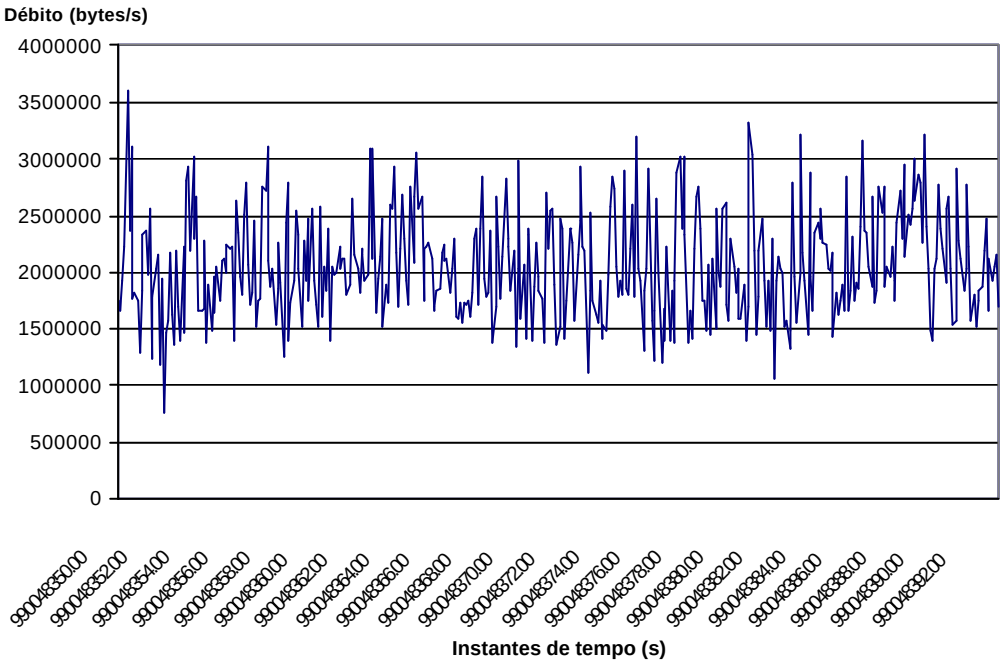

Gráfico 7 - Débito do traço com 100ms de intervalo de contagem

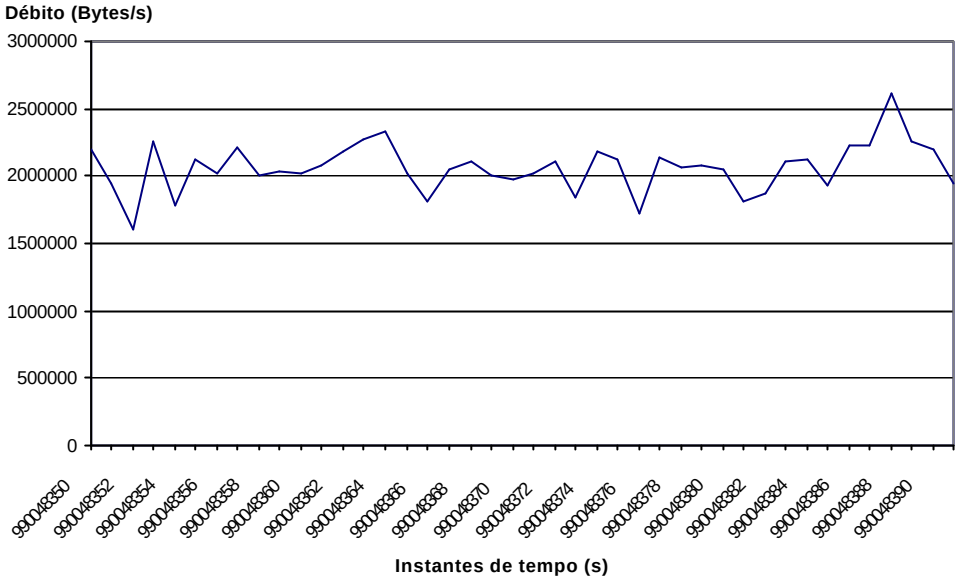

Gráfico 8 - Débito do traço com 1s de intervalo de contagem

Nestes gráficos é possível observar que quanto maior é o intervalo de contagem, mais suave é a forma do gráfico. Este comportamento é o esperado, uma vez que quanto maior é o intervalo de contagem, maior é o número de pacotes contabilizados para cada intervalo, filtrando transições mais abruptas.

No Gráfico 9 podemos ver a contribuição do débito de cada um dos fluxos para o débito total, tendo-se considerado um intervalo de contagem de 1 segundo. Cada fluxo representado neste gráfico corresponde ao fluxo com o mesmo número de ordem, representado nos Gráficos 4 e 5. A partir deste gráfico, pode-se concluir que os fluxos mais importantes (fluxos com maior número de bytes transferidos) são os que mais contribuem para a forma do débito total. Foram considerados 10 fluxos, cujo tráfego corresponde a aproximadamente 50% do tráfego total. Note-se ainda que a duração da maioria dos fluxos individuais considerados corresponde à duração do traço (à excepçãos dos fluxos 3 e 9).

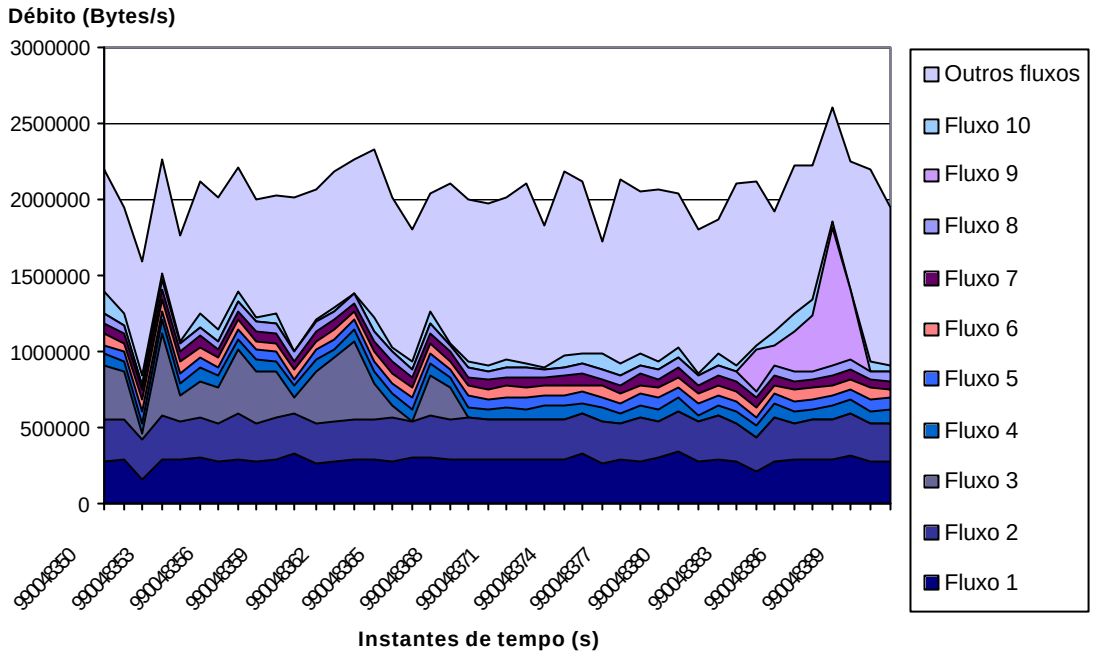

Gráfico 9 - Contribuição dos fluxos para o débito do traço

Resposta 7

No instante 990048367.932311 inicia-se uma ligação HTTP entre o servidor 720908 e o cliente 12517377. São trocados 69 pacotes, sendo 38 entre o servidor e o cliente e os restantes 31 entre o cliente e o servidor. A listagem dos pacotes é apresentada a seguir.

```
1   2 990048367.932311 12517377 720908 6 48 49180 80 02 2439902516 0 32768
2   1 990048367.934002 720908 12517377 6 48 80 49180 12 3652394470 2439902517 32120
3   2 990048368.063948 12517377 720908 6 40 49180 80 10 2439902517 3652394471 32768
4   2 990048368.176132 12517377 720908 6 437 49180 80 18 2439902517 3652394471 32768
5   1 990048368.178181 720908 12517377 6 40 80 49180 10 3652394471 2439902914 31723
6   1 990048368.181335 720908 12517377 6 1500 80 49180 18 3652394471 2439902914 32120
7   1 990048368.181496 720908 12517377 6 1500 80 49180 18 3652395931 2439902914 32120
8   2 990048368.302169 12517377 720908 6 40 49180 80 10 2439902914 3652395931 32768
9   1 990048368.305346 720908 12517377 6 1500 80 49180 18 3652397391 2439902914 32120
10  1 990048368.305671 720908 12517377 6 1500 80 49180 18 3652398851 2439902914 32120
11  2 990048368.350959 12517377 720908 6 40 49180 80 10 2439902914 3652397391 32768
12  1 990048368.354674 720908 12517377 6 1500 80 49180 18 3652400311 2439902914 32120
13  1 990048368.354822 720908 12517377 6 1109 80 49180 18 3652401771 2439902914 32120
14  2 990048368.430491 12517377 720908 6 40 49180 80 10 2439902914 3652400311 32768
```

```
2 990048368.506135 12517377 720908 6 431 49180 80 18 2439902914 3652402840 32768
1 990048368.530615 720908 12517377 6 1500 80 49180 18 3652402840 2439903305 32120
1 990048368.530740 720908 12517377 6 765 80 49180 18 3652404300 2439903305 32120
2 990048368.698898 12517377 720908 6 40 49180 80 10 2439903305 3652405025 32768
2 990048368.704751 12517377 720908 6 434 49180 80 18 2439903305 3652405025 32768
1 990048368.711174 720908 12517377 6 1500 80 49180 18 3652405025 2439903699 32120
1 990048368.711264 720908 12517377 6 416 80 49180 18 3652406485 2439903699 32120
2 990048368.832495 12517377 720908 6 40 49180 80 10 2439903699 3652406485 32768
2 990048368.850317 12517377 720908 6 434 49180 80 18 2439903699 3652406861 32768
1 990048368.856396 720908 12517377 6 1500 80 49180 18 3652406861 2439904093 32120
1 990048368.856690 720908 12517377 6 1500 80 49180 18 3652408321 2439904093 32120
1 990048368.856792 720908 12517377 6 1500 80 49180 18 3652409781 2439904093 32120
1 990048368.856999 720908 12517377 6 1500 80 49180 18 3652411241 2439904093 32120
1 990048368.857112 720908 12517377 6 1500 80 49180 18 3652412701 2439904093 32120
1 990048368.857390 720908 12517377 6 1500 80 49180 18 3652415621 2439904093 32120
1 990048368.857511 720908 12517377 6 1500 80 49180 18 3652417081 2439904093 32120
1 990048368.857636 720908 12517377 6 1500 80 49180 18 3652418541 2439904093 32120
2 990048368.978299 12517377 720908 6 40 49180 80 10 2439904093 3652409781 32768
1 990048368.981508 720908 12517377 6 1500 80 49180 18 3652420001 2439904093 32120
1 990048368.981833 720908 12517377 6 1500 80 49180 18 3652421461 2439904093 32120
1 990048368.981975 720908 12517377 6 1500 80 49180 18 3652422921 2439904093 32120
2 990048368.984740 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048368.984752 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048368.984780 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048368.984786 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
1 990048368.988229 720908 12517377 6 1500 80 49180 18 3652424381 2439904093 32120
1 990048368.988405 720908 12517377 6 1500 80 49180 18 3652425841 2439904093 32120
1 990048368.988518 720908 12517377 6 1500 80 49180 18 3652427301 2439904093 32120
1 990048368.988722 720908 12517377 6 1500 80 49180 18 3652428761 2439904093 32120
1 990048368.988855 720908 12517377 6 1500 80 49180 18 3652414161 2439904093 32120
2 990048369.105050 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048369.106304 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048369.106352 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048369.114212 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048369.114217 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048369.114223 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048369.114302 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
2 990048369.114404 12517377 720908 6 40 49180 80 10 2439904093 3652430221 32768
1 990048369.117500 720908 12517377 6 1500 80 49180 18 3652430221 2439904093 32120
1 990048369.117741 720908 12517377 6 1500 80 49180 18 3652431681 2439904093 32120
1 990048369.117897 720908 12517377 6 1500 80 49180 18 3652433141 2439904093 32120
1 990048369.118032 720908 12517377 6 1500 80 49180 18 3652434601 2439904093 32120
2 990048369.287980 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
1 990048369.291389 720908 12517377 6 1500 80 49180 18 3652437521 2439904093 32120
1 990048369.291479 720908 12517377 6 1500 80 49180 18 3652438981 2439904093 32120
```

```
1 990048369.291674 720908 12517377 6 1500 80 49180 18 3652440441 2439904093 32120
1 990048369.291765 720908 12517377 6 532 80 49180 18 3652441901 2439904093 32120
2 990048369.412466 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
2 990048369.414422 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
2 990048369.414488 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
2 990048369.414507 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
1 990048369.839384 720908 12517377 6 1500 80 49180 18 3652436061 2439904093 32120
2 990048369.959320 12517377 720908 6 40 49180 80 10 2439904093 3652442393 32768
1 990048385.717669 720908 12517377 6 40 80 49180 11 3652442393 2439904093 32120
2 990048385.835470 12517377 720908 6 40 49180 80 10 2439904093 3652442394 32768
```

No Gráfico 10 é feita a representação dos números de sequência dos pacotes, em função do tempo, enviados pelo servidor para o cliente.

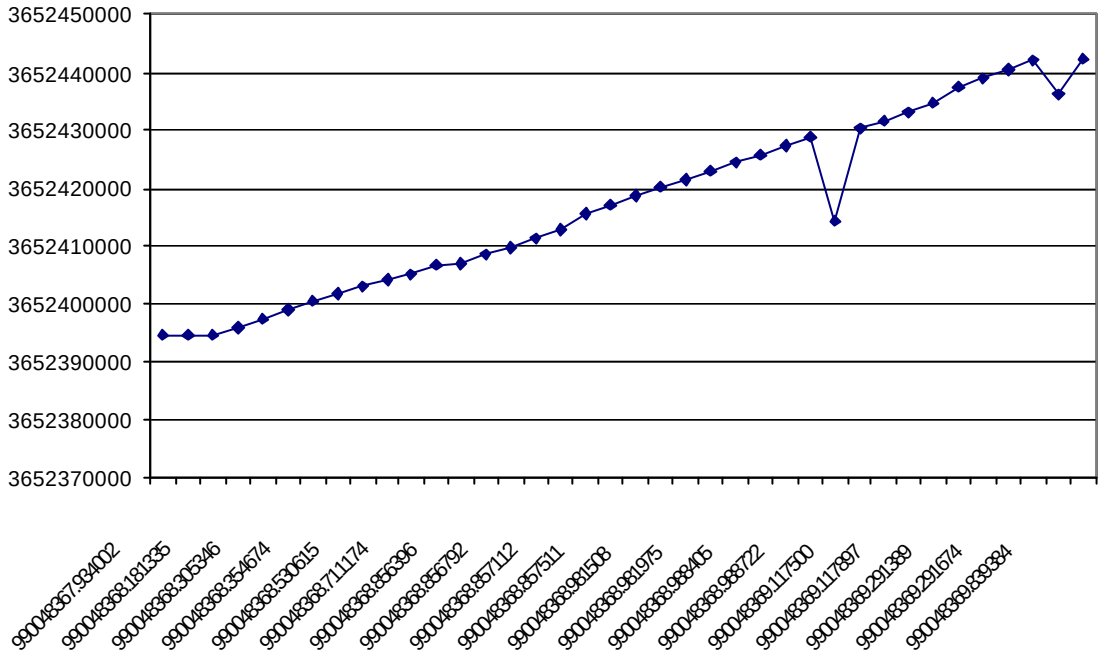

Gráfico 10 - Evolução temporal dos números de sequência dos pacotes enviados peloservidor

O número de sequência cresce de forma aproximadamente linear. As excepções a este funcionamento ocorrem no início (instantes 990048367.934002 a 990048368.181335), em que o número de sequência permanece constante, e nos instantes 990048368.988855 e 990048369.839384 em que o número de sequência diminui apenas para uma transmissão.

Os três primeiros pacotes trocados entre cliente e servidor são relativos ao estabelecimento da conexão TCP através do mecanismo *three-way handshake* (SYN, SYN+ACK, ACK). O quarto pacote trocado é referente ao pedido HTTP, sendo o quinto pacote a confirmação de recepção pelo servidor. No sexto pacote é enviado o primeiro segmento de dados.

```
1  2 990048367.932311 12517377 720908 6 48 49180 80 02 2439902516 0 32768
2  1 990048367.934002 720908 12517377 6 48 80 49180 12 3652394470 2439902517 32120
3  2 990048368.063948 12517377 720908 6 40 49180 80 10 2439902517 3652394471 32768
4  2 990048368.176132 12517377 720908 6 437 49180 80 18 2439902517 3652394471 32768
5  1 990048368.178181 720908 12517377 6 40 80 49180 10 3652394471 2439902914 31723
6  1 990048368.181335 720908 12517377 6 1500 80 49180 18 3652394471 2439902914 32120
```

Como os pacotes enviados pelo servidor durante esta fase da ligação não incluem dados, o número de sequência não cresce – o aumento do número de sequência está directamente relacionado com o número de *bytes* a confirmar.

No instante 990048368.988855 há um decréscimo do número de sequência. Se analisarmos os pacotes antes deste instante de tempo podemos observar que entre o pacote 28 e o pacote 29 falta um pacote. Uma vez que o número de sequência é incrementado pelo número de bytes a confirmar, e sendo a diferença entre os números de sequência dos pacotes 29 e 28 igual a 3652415621 – 3652412701 = 2920 (igual a 1500 bytes – 40 bytes + 1500 bytes – 40 bytes) pode-se concluir que se perdeu um pacote, uma vez que apenas foram transmitidos 1460 bytes de dados. O número de sequência que o receptor estaria à espera seria 3652412701 + 1460 = 3652414161.

```
28 1 990048368.857112 720908 12517377 6 1500 80 49180 18 3652412701 2439904093 32120
29 1 990048368.857390 720908 12517377 6 1500 80 49180 18 3652415621 2439904093 32120
30 1 990048368.857511 720908 12517377 6 1500 80 49180 18 3652417081 2439904093 32120
31 1 990048368.857636 720908 12517377 6 1500 80 49180 18 3652418541 2439904093 32120
32 2 990048368.978299 12517377 720908 6 40 49180 80 10 2439904093 3652409781 32768
33 1 990048368.981508 720908 12517377 6 1500 80 49180 18 3652420001 2439904093 32120
34 1 990048368.981833 720908 12517377 6 1500 80 49180 18 3652421461 2439904093 32120
35 1 990048368.981975 720908 12517377 6 1500 80 49180 18 3652422921 2439904093 32120
36 2 990048368.984740 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
```

```
37 2 990048368.984752 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
38 2 990048368.984780 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
39 2 990048368.984786 12517377 720908 6 40 49180 80 10 2439904093 3652414161 32768
40 1 990048368.988229 720908 12517377 6 1500 80 49180 18 3652424381 2439904093 32120
41 1 990048368.988405 720908 12517377 6 1500 80 49180 18 3652425841 2439904093 32120
42 1 990048368.988518 720908 12517377 6 1500 80 49180 18 3652427301 2439904093 32120
43 1 990048368.988722 720908 12517377 6 1500 80 49180 18 3652428761 2439904093 32120
44 1 990048368.988855 720908 12517377 6 1500 80 49180 18 3652414161 2439904093 32120
```

O cliente, ao verificar que falta um pacote, confirma os pacotes que recebe entretanto com o número de *acknowledge* igual ao número de sequência do pacote que está à espera O servidor ao receber múltiplos *acknowledges* com o mesmo número (pacotes 36-39), faz a retransmissão do pacote em falta (mecanismo de retransmissão rápida).

Um processo semelhante ocorre no instante 990048369.839384 em que é feita a retransmissão de um pacote que foi perdido entre o envio, pelo cliente, dos pacotes 56 e 58.

```
56 1 990048369.118032 720908 12517377 6 1500 80 49180 18 3652434601 2439904093 32120
57 2 990048369.287980 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
58 1 990048369.291389 720908 12517377 6 1500 80 49180 18 3652437521 2439904093 32120
59 1 990048369.291479 720908 12517377 6 1500 80 49180 18 3652438981 2439904093 32120
60 1 990048369.291674 720908 12517377 6 1500 80 49180 18 3652440441 2439904093 32120
61 1 990048369.291765 720908 12517377 6 532 80 49180 18 3652441901 2439904093 32120
62 2 990048369.412466 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
63 2 990048369.414422 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
64 2 990048369.414488 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
65 2 990048369.414507 12517377 720908 6 40 49180 80 10 2439904093 3652436061 32768
66 1 990048369.839384 720908 12517377 6 1500 80 49180 18 3652436061 2439904093 32120
```

Pode ainda observar-se que aproximadamente 16 segundos depois de ser enviado o último pacote pelo servidor para o cliente, o servidor envia um pacote (pacote 68) de finalização da conexão (*flag* FIN activa) ao que o cliente responde com um *acknowledge* (pacote 69). A terminação da ligação foi feita pelo servidor e deve-se provavelmente a um *timeout* da ligação HTTP.

Resposta 8

No Gráfico 11 é representada a função distribuição complementar do tempo entre chegadas de pacotes na interface 1 e a função distribuição complementar da distribuição exponencial.

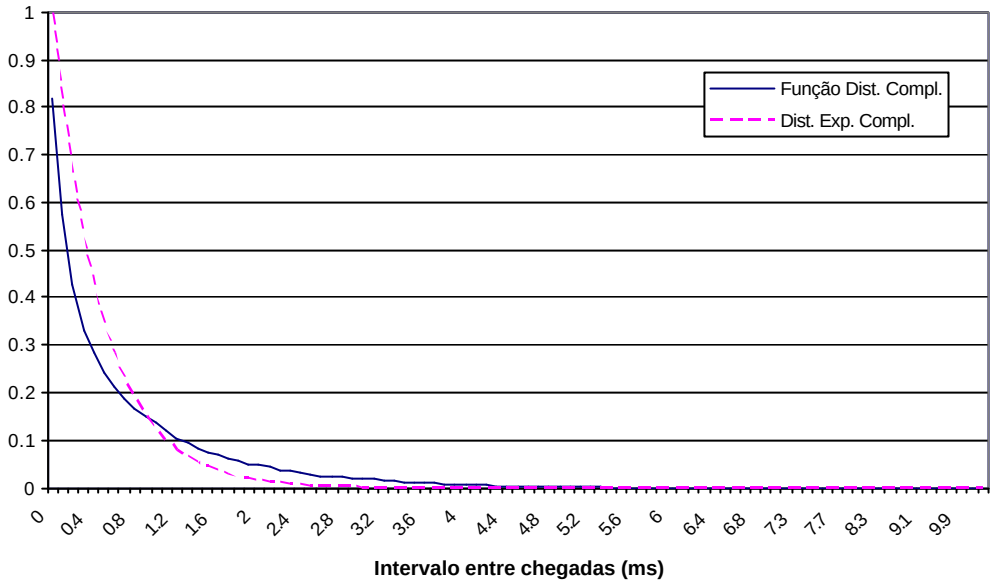

Gráfico 11 - Função distribuição complementar do tempo entre chegadas na interface 1

A distribuição exponencial complementar é definida por $e^{-\lambda x}$, sendo λ a taxa de chegada. Recorrendo ao estimador de máxima verosimilhança demonstra-se que $\lambda = 1 / E(x)$, em que E(x) é a média amostral dos tempos entre chegadas. A partir dos valores obtidos para os tempos entre chegadas obteve-se um valor de λ igual a 2110,13 pacotes/s. Comparando as duas curvas no Gráfico 11 pode afirmar-se que estas são aproximadas, podendo-se concluir que o intervalo de tempo entre chegadas segue aproximadamente uma distribuição exponencial, pelo que o processo de chegada de pacotes é um processo de Poisson.

No Gráfico 12 são representadas as mesmas funções com escala logarítmica nas ordenadas.

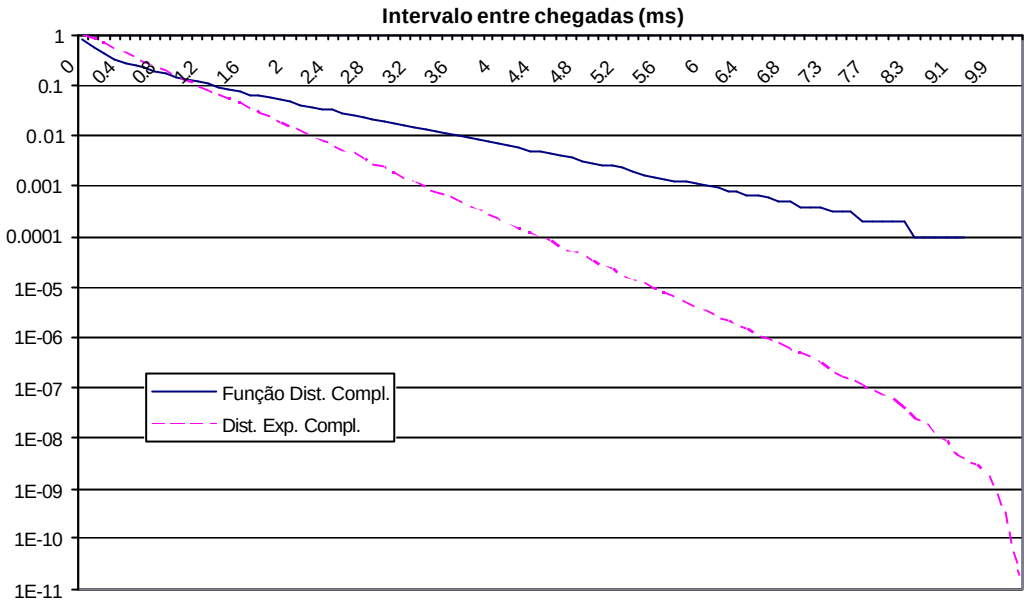

Gráfico 12 - Gráfico 10 com escala logarítmica nas ordenadas