A Packet Arrival Model for Wolfenstein Enemy Territory Online Server Discovery Traffic

Grenville Armitage
Centre for Advanced Internet Architectures.
Swinburne University of Technology
Melbourne, Australia
Email: garmitage@swin.edu.au

Abstract—Clients for online multiplayer first person shooter (FPS) games typically discover game servers through a two-step process. Clients initially query a well-known master server for a list of currently registered game servers, and then sequentially probe each game server in the order they were returned by the master server. The starting and stopping of clients over time creates a 24-hour cycle of ‘background noise’ (probe traffic) impacting on registered game servers, independent of a given server’s actual popularity with players. Based on over 10 million probe packets from two topologically distinct Wolfenstein Enemy Territory servers in 2006, this paper shows that probe arrivals are uncorrelated and exhibit exponentially distributed inter-probe intervals during both busiest and least-busy hours of the 24-hour cycle. A modified Laplace curve is then shown to be a reasonable estimator of $\lambda$ for the exponentially distributed probe arrivals during any hour of the day. The ability to easily synthesise probe traffic patterns will augment existing approaches to modeling the IP traffic loads experienced by game servers and network devices attached to game servers.

I. INTRODUCTION

In the past 6+ years there has been a significant recognition of the popularity and importance of Internet-based multiplayer First Person Shooter (FPS) games (such as Quake III Arena [1], Half-Life Counterstrike [2], Wolfenstein Enemy Territory [3] [4], and Half-Life 2 [5]). FPS games typically operate in a client-server mode, with game servers being hosted by internet service providers (ISPs), dedicated game hosting companies and individual enthusiasts. Although individual FPS game servers typically host from less than 10 to around 30+ players, there are usually many thousands of individually operated game servers active on the Internet at any given time [6]. Due to the fast-paced and highly interactive nature of FPS games, players seek out game servers having predictable latency and low packet loss rates. Consequently, a key challenge for those hosting game servers is to understand the impact on their own Internet connection of actually hosting one or more game servers.

To date most game traffic research has focused on characterising and modeling the network traffic experienced by a game server while people are actually playing the game (for example [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17]). However, FPS game servers also experience a constant ‘background noise’ of server-discovery traffic from (tens of) thousands of clients around the planet [18]. Server discovery operates similarly for many types of FPS games (as a consequence of the decentralised nature of game server hosting). First, a game client queries a master server unique to the particular game (pre-configured into the game client software). The master server returns a list of hundreds or thousands of IP addresses and port numbers representing game servers registered as currently active. The client then steps through this list, probing each listed game server for information about map type, game type and number of players (typically a brief UDP packet exchange). This probe process also estimates the client to server RTT at the time of the probe. All this information is presented to the player as it is gathered, who then selects a game server to join.

A given client will send out hundreds or thousands of probe packets to find and join only one game server. Consequently, individual game servers end up receiving, and responding to, tens of thousands of probe packets unrelated to the people actually playing at any given time. The background noise due to probe traffic fluctuates over time as clients around the Internet startup and shutdown.

This paper focuses on the impact, from a game server’s perspective, of hundreds or thousands of (largely) unrelated game clients independently probing all available game servers, 24 hours a day. We model the general properties of such probe traffic based on the experience of two topologically distinct Wolfenstein Enemy Territory (ET) servers in 2006. (ET utilises the Quake III Arena game engine, and inherits a similar server discovery mechanism.) Over 10 million individual probe packets were captured during January, April and July 2006. Consistent with previous work [18], probe arrivals were seen to fluctuate on a 24-hour cycle and be independent of a particular game server’s popularity with players. This paper’s more finely-grained analysis reveals that probe traffic shows exponentially distributed inter-probe intervals during both busiest and least-busy hours of the cycle. Ultimately this knowledge will assist in the creation and refinement of synthetic models of the IP traffic loads experienced by routers connecting game servers to the internet (and by systems actually hosting game servers).

The rest of this paper is organised as follows. Section II describes the specific discovery process used by ET, and summarises the data gathered from our two ET servers. Section III summarises previous related work on server discovery probe traffic. Section IV discusses the observed probe traffic patterns,
and Section V discusses some implications, limitations and areas for future work. Section VI concludes the paper.

II. COLLECTING CLIENT PROBE TRAFFIC

A. The Enemy Territory client’s probe sequence

A public ET game server will automatically register itself with the ET master server at etmaster.idsoftware.com. This master server becomes a rendezvous point for clients around the planet who wish to know what servers are available at any point in time.

B. Collecting real-world probe traffic from two ET servers

During early 2006 two identically-configured ET game servers were monitored 24 hours a day, 7 days a week in Australia. One server was located in Melbourne, Australia (at the Centre for Advanced Internet Architecture, CAIA) and the other in Canberra, Australia (hosted by Grangenet [20]). Both were registered with the ET master server for public play. Students from CAIA tended to play on the Melbourne server, often triggering the arrival of additional players and causing long periods of active game-play. The Grangenet server was largely left alone, and saw far less active game-play. All packets entering and leaving both servers were captured for subsequent analysis.

Figure 1 illustrates each stage of an ET client’s server discovery process. The client sends a short UDP request packet to the master server on port 27950 (step 1), eliciting one or more UDP packets that contain all the currently registered game servers (step 2). As the list is retrieved (step 3), the game client begins probing each game server in the order in which they were listed by the master server (step 4). The game client populates its on-screen server browser (step 5) as game servers respond with their current status information. At any time during (or after) step 5 the player may chose a specific server, from which the client chooses server to join from Game Browser (step 6).

In early 2006 the ET master server’s response contained between 2800 and 3100 servers spread across 26 to 28 UDP packets of up to 810 bytes each (6 bytes of IPv4 address and UDP port number for each registered server) [19]. Client probes are 43-byte UDP/IP packets with payload starting 0xFFFFFFFF followed by the ASCII text “getinfo xxx”. An active game server will respond promptly with a multi-hundred byte UDP packet beginning with 0xFFFFFFFF and the text “infoResponse”.

As a small optimisation, an ET client partially overlaps steps 3 and 4. The first 16 game servers are probed in sequence as soon as the first reply packet arrives from the master server, with additional probes sent as previous probes are answered. No more than 16 probes remain outstanding (unanswered) at any one time (although the actual number varies depending on how long it takes previously probed servers to reply). An ET client would take roughly one minute to probe an average of 3000 servers [19].

Table I shows the total number of probe packets received by each server in January, April and July 2006.

<table>
<thead>
<tr>
<th>Number of countries seen</th>
<th>Jan 2006</th>
<th>Apr 2006</th>
<th>Jul 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>5219842</td>
<td>2740068</td>
<td>2599319</td>
</tr>
<tr>
<td>2nd</td>
<td>3835542</td>
<td>2835542</td>
<td>2492247</td>
</tr>
<tr>
<td>3rd</td>
<td>3200000</td>
<td>235594</td>
<td>263581</td>
</tr>
<tr>
<td>4th</td>
<td>250008</td>
<td>223691</td>
<td>183806</td>
</tr>
<tr>
<td>5th</td>
<td>177775</td>
<td>190035</td>
<td>144541</td>
</tr>
</tbody>
</table>

Table II shows the top five countries probing the CAIA server each month.

<table>
<thead>
<tr>
<th>Top 5 countries probing CAIA server each month</th>
</tr>
</thead>
<tbody>
<tr>
<td>---------</td>
</tr>
<tr>
<td>NL</td>
</tr>
<tr>
<td>US</td>
</tr>
<tr>
<td>FR</td>
</tr>
<tr>
<td>AU</td>
</tr>
<tr>
<td>CA</td>
</tr>
</tbody>
</table>

Identifying the geographic source of probe traffic was performed using MaxMind’s free GeoLite Country database [21]. Maxmind claim their GeoLite Country database correctly maps 97% of all IP addresses to country codes, and provides significant coverage of the active IP address space (of the 8.26M probes seen across all three months at the CAIA ET server, only 2353 could not be resolved to a country). Table II ranks the top five countries probing the CAIA server (and the number of probes sent) in January, April and July 2006.

Authorized licensed use limited to: SWINBURNE UNIV OF TECHNOLOGY. Downloaded on January 4, 2010 at 22:05 from IEEE Xplore. Restrictions apply.
Both game servers were connected to the Internet over 100Mbit/sec LAN interfaces and relatively uncongested academic research links. Even pessimistically assuming 10Mbit/sec of available capacity to the wider Internet, serialisation delays for the 43-byte probe packets would be less than 50 microseconds. It seems reasonable to believe that the inter-probe distributions discussed in the following sections were not significantly skewed by link-layer issues close to the game servers.

III. RELATED WORK

It does not appear that any prior work has explored and demonstrated uncorrelated, exponentially distributed inter-packet arrival times for server discovery probe traffic. The ET servers used in this paper were the basis for an earlier study of the relationship between game-play (clients connected and playing) and probe traffic impacting on public game servers. Zander, et al, observed that game-play and probe traffic operate on unrelated 24-hour cycles [18]. Probe traffic would rise and fall as the number of potential players changed over time, whereas game-play traffic depended only on the number of actual people who chose to play on our servers. (Few people played on the Grangenet server, yet it saw very similar levels of probe traffic to that seen by the CAIA server). Traffic from different geographical regions was also shown to have distinct, human-driven 24-hour cycles - both probe and game-play activity would peak during each region’s late afternoon and evening. However, [18]’s use of aggregated per-flow statistics prevented detailed analysis of inter-probe intervals over time.

In [19] the ET server discovery process was analysed from the client’s perspective. Every 36 minutes the ET master server would vary the location of every game server within the list returned in step 2 of Figure 1. Consequently, over long periods of time, a given game server would find itself 1st, 2nd.... Nth in the list returned by the master server. We used this information to propose some client-side optimisations to the server probe algorithm in Figure 1.

Distantly related work [9] has focused on dynamic server re-discovery, redirecting players from one game server to a closer game server based on inferring geographic locality from client IP addresses. The characteristics of server-discovery traffic itself was not addressed.

The present paper may be considered a successor to [18].

IV. EVALUATION OF OBSERVED PROBE TRAFFIC

A. The human origins of most probe traffic

Figure 2 shows the average number of probes seen per hour over a typical 24-hour period from all source addresses. CAIA server statistics are shown for the months of January, April and July 2006. The Grangenet server’s statistics are shown for April 2006. The x-axis is in terms of ‘hours relative to GMT+10:00’, where 0 is midnight and 23 is 11pm in the GMT+10:00 timezone (Eastern Standard Time for both ET servers).

Focusing on the CAIA server, Figure 3 shows the average number of probes per hour of the calendar week during April 2006 from the United States (US) and Poland (PL). The 24-hour cycle is quite evident, as is a distinct phase difference between PL and US distributions. The x-axis is in terms of ‘hours relative to GMT+10:00’, where 0 is midnight on Sunday and 167 is 11pm the following Saturday in the GMT+10:00 timezone.

Figures 2 and 3 strongly suggest that server-discovery probe traffic is largely a human-initiated process. The aggregate traffic seen in Figure 2 is made up of regionally-specific probe traffic, which Figure 3 shows is strongly influenced by the times of day that potential players are available from each geographic region. The offset of Figure 2’s January 2006 curve may also be attributed to human factors behind the probe traffic. Most of April 2006 and all of July 2006 were ‘summer time’ in the northern hemisphere (from where the majority of probes originate), with increasingly longer days and shorter nights. January 2006 was ‘standard time’ (and the middle
of winter). Not surprisingly the peak probing hour (relative to GMT+10:00) has shifted roughly 1 to 2 hours between wintertime and summertime because of ‘daylight saving time’ and players simply modifying their daily patterns of life. Finally, Figure 2 shows that both the CAIA and Grangenet server saw almost identical levels of probe traffic, independent of the CAIA server’s popularity with players.

B. Exponentially distributed inter-arrival times

The previous section largely validates insights first seen in [18]. The current paper’s main contribution is identifying and demonstrating that inter-probe intervals can be closely modelled by an exponential distribution at small time scales.

Figures 4 and 5 show the cumulative distribution of inter-probe intervals for the busiest and least busy hours of the day respectively across April 2006. (The busiest and least-busy hours of a typical day were identified from the data used to create Figure 2. Then the actual inter-probe intervals from each of those hours in every day of April was utilised to create Figures 4 and 5.)

For comparison the following exponential curve is also plotted:

\[ CDF = 1 - e^{-\lambda x} \]  

(1)

where \( \lambda \) is derived from a single characteristic of the experimentally acquired data:

\[ \lambda = \frac{\log(2)}{\text{MedianIntv}} \]  

(2)

where MedianIntv is the median inter-probe interval during the hours of interest. (\( \lambda \) is derived from the median, rather than directly from the mean, to minimise influence from a handful of unrepresentative outliers in the raw inter-probe interval dataset.)

C. Uncorrelated inter-arrival times

Figure 7 shows the auto-correlation of 9962 inter-probe intervals (measured during the busiest hour of April 2006) for ‘lag’ between 0 and 100. The clear peak at \( \chi = 0 \) and low values for \( \chi \geq 1 \) shows that individual inter-probe intervals are uncorrelated over periods of at least tens of seconds. (This seems reasonable - the human players initiating FPS server-discovery events are acting largely independently of each other.) Consequently, getinfo probe arrivals over short periods (less than an hour) may reasonably be simulated by a random process that produces independent, exponentially distributed values.
Fig. 7. Auto-correlation of inter-probe intervals, busiest hour April 2006

**D. Hourly variation in median inter-probe intervals**

It is clear from Figures 4 and 5 that the median inter-probe interval itself fluctuates over time. Therefore any synthesis of server-discovery probe traffic over a multiple hours or days requires a model to predict the variation in median inter-probe interval for each hour of the day.

A variation on the Laplace distribution makes a reasonable first approximation. Figure 8 shows the average median inter-probe interval per hour over 24 hours for the CAIA server in January and April. Compared to the almost-sinusoidal variation in Figure 2, Figure 8 is dominated by a daily `peak'.

For comparison, the following Laplace distribution has been overlaid on each month’s actual distribution (with ``(laplace)'' appended to its label):

\[
\text{MedianIntv} = \text{Base} + (\text{Peak} - \text{Base}) \times (e^{-|x/4|}) \quad (3)
\]

*Base* and *Peak* are the shortest and longest per-hour median inter-probe intervals (respectively) seen across the 24-hour period. We vary \(x\) from -12 to 11, and then wrap the curve onto a 24-hr period such that the peak (normally at \(x=0\)) occurs on the hour at which the actual distribution peaks.

Consequently, by measuring the median inter-probe interval during the busiest and least-busy hours of the day, and noting which actual hour is least busy, we can establish the *Base* and *Peak* parameters for equation 3. Equation 3 may then be used in equation 2 to derive an approximate \(\lambda\) to use in equation 1 for arbitrary hours of the day.

**E. An undercurrent of automated probing**

Although server-discovery probe traffic is primarily driven by human activity, Figure 3 reveals an interesting fact - not every country 'goes to sleep'. Whilst probe traffic from Poland drops almost to zero once every 24 hours, the United States originates some finite level of probe traffic at all hours of the day. This cannot simply be attributed to the US spanning multiple time-zones, as the west coast and east coast both share a number of hours in the early morning where people would normally be asleep.

Closer inspection of the raw traffic reveals that a very small number of probe sources are automated systems, regularly and continuously polling game servers all around the world. These systems (such as ServerSpy [22]) are designed to create rankings of players across different genres of FPS games, and across all public servers for particular FPS genres. It is likely that such services contribute to the non-zero base level of US traffic in Figure 3.

The existence of automated probe sources may be of some interest if we are attempting to model precisely the probe-by-probe arrival of packets on a particular server. But when dealing with longer-term traffic statistics over tens of minutes or hours, the dominant influence comes from probes triggered by human activity.

**F. Trends over multiple months**

It was noted in [18] that the peak probe activity is likely to be proportional to the density of potential players around the planet. The descending peaks of the January, April and July 2006 CAIA curves in Figure 2 reveals that the total number of ET clients being turned on and off each day was slowly declining during the first half of 2006. (This does not reflect a change in the CAIA game server's own popularity. Had the number of potential ET players been increasing, we would have seen an increase in probe traffic over time.) The sources of long-term fluctuations in peak probe traffic are beyond the scope of this paper. Nevertheless, any detailed synthesis of probe traffic over multiple calendar months would do well to consider such trends.

**V. IMPLICATIONS, LIMITATIONS AND FUTURE WORK**

**A. The impact on individual game servers**

This paper clarifies the reality that distributed FPS server discovery traffic on a particular game server scales with the number of potential players. It does not scale with the number
of potential game servers, a game server’s own popularity or the game server operator’s attempts to limit the number of concurrent clients who can actually play at any given time. If the particular genre of game became orders of magnitude more popular the increase in potential players would result in orders of magnitude more ‘background noise’ per unit time for each and every public game server.

The relevance becomes clearer when we consider the consequent outbound traffic. Every ET getinfo xxx probe packet triggers a 317-byte infoResponse reply from a targeted game server (varying +/- a few bytes over time). Reply packets inherit the same inter-packet interval distribution as the inbound probe packets, yet consume far more link capacity due to their size. For example, in January 2006 the CAIA server saw 132Mbytes of getinfo xxx probes and replied with 973Mbytes of infoResponse traffic - over 1Gbyte of aggregate traffic to satisfy the curiosities of potential players who are unlikely to ever play on the CAIA server. Every other public ET server would have also experienced this monthly server discovery traffic load. Furthermore, the transmission times of infoResponse packets will inherit the uncorrelated and exponential distribution of the inbound getinfo requests.

B. Limitations and future work

This paper illustrates its insights using probe traffic impacting on two Wolfenstein Enemy Territory servers during the first half of 2006. However, the absolute numbers (peak probes per hour, shortest median inter-probe intervals, etc) will differ for other FPS games (such as other Quake variants or Half-Life 2 variants), influenced by the different densities of active game clients for particular FPS game types. For example, ET was particularly dominated by European players. The peaks and troughs in Figure 2 (and hence in Figure 8) would flatten out if the density of potential players was spread more evenly around the planet’s timezones. Equation 3 might therefore need to be revised for other FPS games before being used to drive equations 1 and 2.

Future work should include estimating the number of potential players from the probe traffic density over time (for ET and similar FPS games), optimising FPS server discovery to support orders of magnitudes more potential players, and estimating the impact of client-side server discovery optimisations (such as [19]) on daily probe traffic patterns.

VI. CONCLUSION

Over 10 million server discovery probe packet arrivals at two separate Wolfenstein Enemy Territory (ET) servers during the first half of 2006 have been analysed. Probe traffic density is shown to cycle over 24-hour periods, driven largely by humans seeking out public game servers. Probe traffic arrivals are shown to be uncorrelated over tens of seconds, and exponentially distributed during both the busiest and least-busy hours of the day. Given the busiest and least-busy hour’s median inter-probe intervals, a modified Laplace curve is a reasonable estimator of $\lambda$ with which to synthesise exponentially distributed probe arrivals during any hour of the day. As ET game servers respond immediately to individual getinfo probes these results predict that the transmission times of outbound (and much larger) infoResponse packets will be similarly uncorrelated and exponentially distributed. This paper thus provides an improved model for simulating traffic due to server-discovery getinfo probe activity on network links connecting to public ET game servers.

VII. ACKNOWLEDGEMENTS

I would like to thank Philip Branch for his valuable encouragement regarding the statistical importance of exponentially distributed packet arrivals, and Sebastian Zander for managing the CAIA and Grangenet game servers from 2004 to 2006.

REFERENCES