An Exponential Arrival Model for Half-Life 2 Server Discovery Traffic

Grenville Armitage, Philip Branch
Centre for Advanced Internet Architectures, Technical Report 090225A
Swinburne University of Technology
Melbourne, Australia
garmitage@swin.edu.au, pbranch@swin.edu.au

***Abstract*—Online multiplayer first person shooter (FPS) clients typically discover public game servers through a two-step process. A master server is queried for a list of current game servers, who are then sequentially probed (creating a 24-hour cycle of *background noise* for FPS game servers). Using 3 million probe packets received by a Half-Life 2 server in February and April 2006, we show arrivals exhibit exponentially distributed, uncorrelated inter-probe intervals during both busiest and least-busy hours of a day. We further show a modified Laplace curve is a reasonable per-hour estimator of λ for the exponentially distributed probe arrivals.**

I. Introduction

First Person Shooter (FPS) online games are increasingly popular and demanding of good IP layer quality of service [1]. Typically operating in client-server mode, game servers are hosted by internet service providers, game hosting companies and individual enthusiasts. Thousands of game servers are active on the Internet at any given time, each one typically allowing between 4 to 30+ players. The fast-pace and interactivity of FPS games drives players to game servers having consistent and low latency. To ensure quality of service, a key challenge is characterising the impact on one's internet connection of hosting a game server.

Most game traffic research has focused on characterising and modelling the network traffic experienced by a game server while people are *actually playing* the game [1], rather than the *background noise* of server-discovery traffic from thousands of non-playing game clients around the planet [2]. This report[1] describes a model for the daily background noise experienced by Half-Life 2 'death match' (HL2DM) game servers. Using 3 million probes received by a HL2DM game server over two months in 2006 we demonstrate that an exponential model reasonably approximates the distribution of probe packet arrivals. This knowledge supplements existing models for simulating IP traffic in and out of game servers. We believe our work to be the first detailed analysis of inter-packet arrival time characteristics for HL2DM server-discovery probe traffic.

HL2DM server-discovery is based on Valve's Steam [3] online authentication system. Individual HL2DM game servers register themselves with the Steam master server at *hlmaster1.hlauth.net*. Within minutes a registered game server will start receiving UDP probe packets from around the Internet, as Steam-based clients begin to 'see' it in the Steam master server's list of active game servers. Probes are manually initiated by potential players through their client's game server browser. They retrieve a list of registered game servers from the Steam master server and probe each game server in sequence - retrieving server state information in multi-hundred byte reply packets. Initial client probes are 53-byte UDP/IP packets containing the ASCII string "*TSource Engine Query*" (TSEQ). A given client will send out hundreds or thousands of *TSEQ* probe packets before joining only one game server. Individual game servers receive, and respond to, tens of thousands of probes packets unrelated to the people actually playing at any given time.

II. Collecting Client Probe Traffic

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

We operated a public HL2DM game server in Melbourne, Australia, continuously during February and April 2006 on *gs.caia.swin.edu.au:27016*, capturing all ingress and egress packets. We saw 1638370 and 1408110 *TSEQ* probes in February and April 2006 respectively. Using MaxMind's GeoLite Country database [4] (claimed to correctly map 97% of all IP

[1]This technical report belatedly archives a paper submitted to IEEE Communications Letters on June 20th 2007. The paper was rejected.

addresses to country codes) we identified multiple countries probing throughout every day. In February 2006 the top three probe sources were the US, Germany and UK (with 512057, 229474 and 123316 probes respectively). In April the top three sources were the US, Germany and Australia (with 402032, 212116 and 126017 probes respectively). Less than 1000 probes could not be resolved to a country code.

Our server's uncongested Internet link ensured measured intervals between *TSEQ* packets were dominated by human-triggered server-discovery events (rather than link-layer queueing or serialisation jitter).

III. Evaluation of observed probe traffic

A. *The human origins of most probe traffic*

Fig. 1 shows the average number of probes per hour of the calendar week during April 2006 from the United States (US), Germany (DE) and Australia (AU). (The x-axis is relative to GMT+10:00, our game server's local timezone. Midnight on Sunday is 0, 11pm the following Saturday is 167.) The 24-hour cycles and distinct phase differences suggest that probe traffic is driven by region-specific human activity.

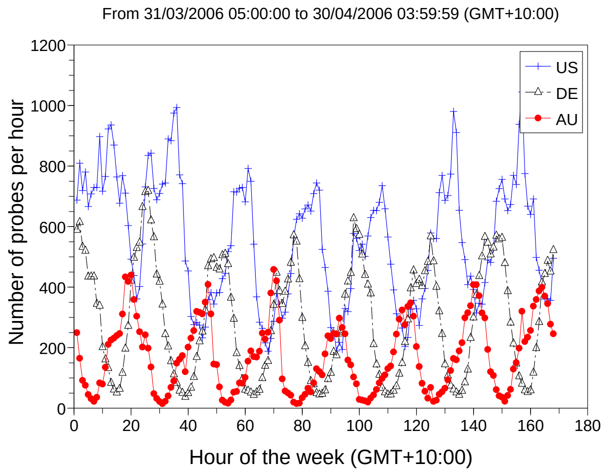

Fig. 1. Probe packets per hour of a week from USA, Germany and Australia

B. *Exponentially distributed inter-arrival times*

Fig. 2 shows the cumulative distribution of inter-probe intervals for the busiest hours of each day during April 2006. For comparison the following exponential curve is also plotted:

$$CDF = 1 - e^{-\lambda x} \tag{1}$$

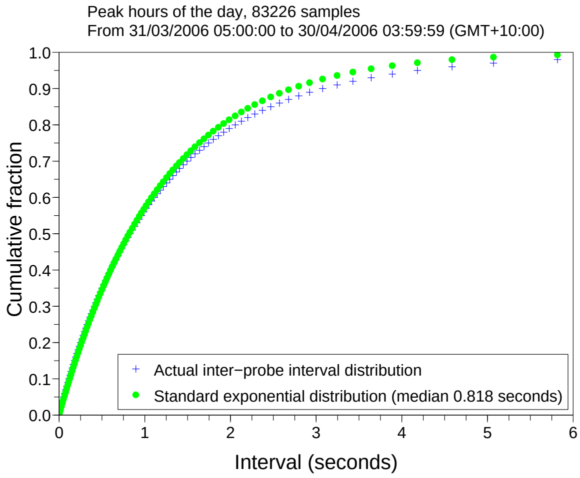

Fig. 2. Distribution of inter-probe intervals during busiest hours of a day

where λ is derived from *MedianIntv*, the median inter-probe interval measured from the experimentally acquired data during the hours of interest:

$$\lambda = \frac{log(2)}{MedianIntv} \tag{2}$$

A similarly close match exists during least-busy hours (excluded for brevity), despite the median inter-probe interval moving out from 0.818 seconds to 2.019 seconds. Q-Q plots of inter-probe intervals vs the synthesised exponential distributions (also excluded for brevity) showed that, for both busiest and least-busy hours, probe traffic arrivals are well modelled by an exponential distribution.

C. *Uncorrelated inter-arrival times*

Fig. 3 shows the auto-correlation of inter-probe intervals measured during the busiest hour of April 2006 for 'lag' between 0 and 100. The peak at x = 0 and low values for $1 \leq x \leq 100$ show individual inter-probe intervals are uncorrelated over periods of at least 80 seconds (consistent with human players initiating independent FPS server-discovery events.) It seems reasonable that TSEQ probe arrivals may be simulated by a random process that produces independent, exponentially distributed values.

D. *Modelling the median inter-probe interval over 24 hours*

Next we model the daily variation in median inter-probe interval. Fig. 4 shows the 24-hour cycle of average median inter-probe interval per hour for our February

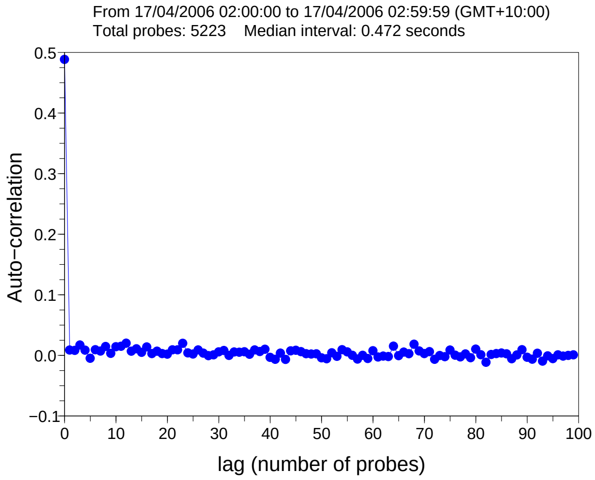

Fig. 3. Auto-correlation of inter-probe intervals, busiest hour April 2006

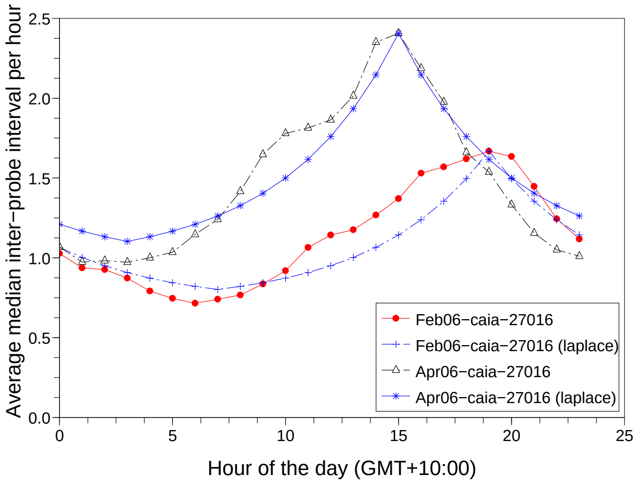

Fig. 4. Median inter-probe interval per hour of a day

and April probe traffic. A modified Laplace curve (labelled "(laplace)") is overlaid, making a reasonable first approximation to the measured distribution.

For any given 24-hour period the modified Laplace curve is given by:

$$MedianIntv = Base + (Peak - Base) * (e^{-|(x/4)|}) \quad (3)$$

where *Base* and *Peak* are the shortest and longest measured per-hour median inter-probe intervals (respectively). We vary x from -12 to 11, and wrap the curve such that the peak (normally at x=0) occurs on the hour at which the measured distribution peaks.

Consequently, from the median inter-probe interval during the busiest and least-busy hours of the day, and the actual hour that is least busy, we establish *Base* and *Peak* for equation 3. Equation 3 may then be used in equation 2 to derive an approximate λ to use in equation 1 for arbitrary hours of the day.

IV. Conclusion

We captured 3 million probe packets to derive an exponential model for probe packet arrivals at a Half-Life 2 server during February and April 2006. This model applies during both the busiest and least-busy hours of a day, and probe arrivals are shown to be uncorrelated over short (tens of seconds) periods of time. Given the busiest and least-busy hour's median inter-probe intervals, a modified Laplace curve is shown to be a reasonable estimator of λ for the exponentially distributed probe arrivals during any hour of the day. As many other FPS games utilise similar, player-triggered server discovery it seems plausible that an exponential arrival model will apply more generally. However, this is subject of additional work.

References

[1] G. Armitage, M. Claypool, and P. Branch, *Networking and Online Games - Understanding and Engineering Multiplayer Internet Games*. United Kingdom: John Wiley & Sons, April 2006.

[2] S. Zander, D. Kennedy, and G. Armitage, "Dissecting server-discovery traffic patterns generated by multiplayer first person shooter games," in *Proceedings of ACM Networks and System Support for Games (NetGames) Workshop*, October 2005.

[3] Valve Corporation, *Welcome to Steam*, http://www.steampowered.com/, as of April 27th 2007.

[4] MaxMind, *GeoLite Country*, http://www.maxmind.com/app/geoip_country, as of April 29th 2007.