An Experimental Estimation of Latency Sensitivity In Multiplayer Quake 3

Grenville J. Armitage

Centre for Advanced Internet Architectures. Technical Report 030405A
Swinburne University of Technology
Melbourne, Australia
garmitage@swin.edu.au

Abstract- **Quantifying the latency sensitivity of potential customers/players is critical for Internet-based game providers when planning the network placement of their game servers. In early 2001 we placed two Quake 3 servers at different locations on the Internet, and instrumented them to gather median latency information on every player who played over a multi-month period. Comparison of server logfiles showed an active yet distinct player population on each server, and the median latency distributions suggest players actively prefer Quake 3 servers less than 150 to 180 milliseconds from the player's location. Quake 3 is often played as a multiplayer, Internet-based, highly interactive "first person shooter" game. Although Quake 3 is nowhere near as popular as games such as Half-life:CounterStrike, we believe our results provide a useful ball-park indicator of latency sensitivity for this class of highly interactive online games.**

Keywords- Online, Games, Latency, Internet, Quake 3

I. INTRODUCTION

It is easy to qualitatively assert that network latency is important for players of highly interactive, online multiplayer games. It is not so easy to quantify this latency sensitivity in a way that is useful to online game service providers who are trying to adequately provision their service. This paper contributes some quantitative bounds to such discussions[1].

In early 2001 we placed two Quake 3 [1] servers at different locations on the Internet, and instrumented them to gather median latency information on every player who played over a multi-month period. Comparison of server logfiles showed an active yet distinct player population on each server, and the median latency distributions suggest players actively prefer Quake 3 servers less than 150 to 180 milliseconds from the player's location. Quake 3 is often played as a multiplayer, Internet-based, highly interactive "first person shooter" game. Although Quake 3 is nowhere near as popular as games such as Half-life:CounterStrike [2], we believe our results provide a useful ball-park indicator of latency sensitivity for this class of highly interactive online games.

---

[1] A version of this paper will be presented at the 11th IEEE International Conference on Networks (ICON 2003) in Sydney, Australia, September 2003

II. RELATED WORK

To the best of our knowledge there has been no other published work on the latency sensitivity of Quake 3 games. A preliminary latency sensitivity study of Half-Life in 2001 concluded that network latency was less important than vagaries in the latencies induced by application/host behaviors and the relative delays between players on the same server [3]. A more recent 2002 study of latency sensitivity for Warcraft3 found that players could tolerate substantial fractions of a second in network latency [4]. Other work in the area of online games has tended to focus on modeling of network and user behaviors [5][6][7] rather than latency sensitivity per se.

III. METHODOLOGY

Key to this study is the use of two distinct Quake 3 servers that appeared (to potential players) as essentially identical except for their apparent latency. The servers themselves were placed at quite different locations on the Internet - Palo Alto, USA and London, UK (roughly 147 milliseconds apart) - in order to ensure a decent spread in the pool of players attracted to each server. Each server then kept a log of every player's IP address and joining, playing, and leaving times. While playing, the server's perception of latency ('ping') time to each client was logged every time the client's player killed another player, was killed by another player, or picked up an object in the game.

A. Apparently Identical Servers

To understand how we made the servers appear identical it is important to understand the two key methods by which players locate and select Quake 3 game servers. First, all public Quake 3 servers register their existence with master servers run by idSoftware (the developers of Quake3). Players then use either the in-game server selection tool or a 3rd party application such as GameSpy3D [8] to search the list of current servers.

Players are presented with only limited information about each server , e.g. the server's name (arbitrarily assigned by the server administrator and unrelated to the server's IP address), the number of players on the server , additional descriptive information provided by the server administrator and a current estimate of the 'ping' time to that server. The potential player is provided with no method for evaluating the topoogical significance of

any game server's IP address (which, aside from the latency, is the only other piece of information by which the two servers would look different).

We arranged for our two servers to offer the exact same map sequence, map time limits, frag (kill) limits, player number limits, and almost identical (yet relatively meaningless) server names - "BLRSV#2 DM 6 map cycle" and "BLRSV#3 DM 6 map cycle". Both servers had an additional server variable (visible with GameSpy3D) claiming the server was located in "Palo Alto, California". We also ran 2 'bots' (computer-generated players) on each server to attract players. To minimise any influence on player choice the bots had the same names on each server.

In the end, potential players were left with apparent current latency (as reported by their client or GameSpy3D application) as the only meaningful differentiator between our two servers.

B.   The Actual Servers

The Quake 3 servers were located 147 milliseconds apart in Palo Alto, Calfornia (USA) and London (UK) (Figure 1). Both servers were configured as follows:

- ◆ Linux Quake3 server verison 1.17 (current at the time)
- ◆ The same 6 standard maps, 20 minutes per map
- ◆ Upto 6 remote players + 2 permanent 'bots'

Palo Alto's server was a rack-mounted 600MHz Celeron with 128MB of RAM running FreeBSD 4.2 and a direct T1 connect to PAIX (the Palo Alto Internet eXchange). London's server was a 900MHz Athlon with 128MB running Linux kernel 2.4.2 and a 10Mb link to the UK academic network.
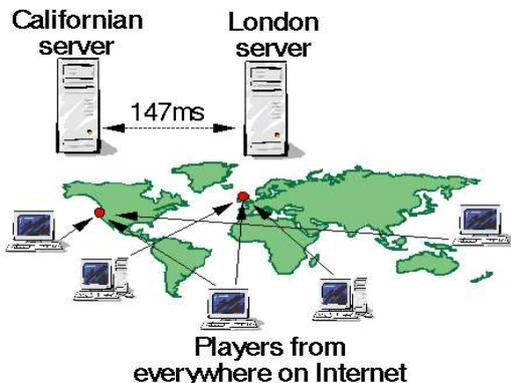
Figure 1 Placing the servers 147milliseconds apart ensured a topologically spread player population

C.   Server Statistics

The Californian server ran from May 17 to August 18, 2001 and saw 5290 unique clients who accumulated a total of 164 'days played'. 338 clients accumulated more than 2 hours total playing time during this period.

The London server ran from May 29 to September 12, 2001 and saw 4232 unique clients who accumulated a total of 77 'days played'. 131 clients accumulated more than 2 hours total playing time during this period.

Figure 2 Cumulative played time on Californian server

Our servers were frequented reasonably consistently across the trial period. Figure 2 and Figure 3 show how the 'days played' (the cumulative time played by every player) increased consistently over the period of our test (aside from three short server outages in London).

D.   Uniqueness of clients

Although every client has an ASCII playername, they have no guarantee of uniqueness. Client IP addresses are also subject to variations within a range if a player's ISP assigns different IP addresses each time the client dials-up each day (so counting by IP address alone could result in over-counting unique clients). Thus, a unique client is identified by the playername and the non-hostname part of the domain name resolved by reverse lookup of the client's IP address. (Clients who connected as the "unknown player" default playername were excluded from our analysis.)

Figure 3 Cumulative played time on London server

E.   Collecting Ping Times

The server's perception of latency ('ping') time to each client was logged every time the client's player killed another player, was killed by another player, or picked up an object in the game. Clients who accumulated less than 10 ping samples per game were removed from the analysis.

IV. RESULTS

The crux of our results is the cumulative distribution of median ping times shown in Figure 4, which reflects

each player's median ping counted at the end of each game (in other words, 'median ping per player per game').

We believe that a player returning to play again is correlated with their satisfaction with their game experience, so Figure 4 captures the rate at which increased latency sours the player's game experience.

Cumulative Median Ping
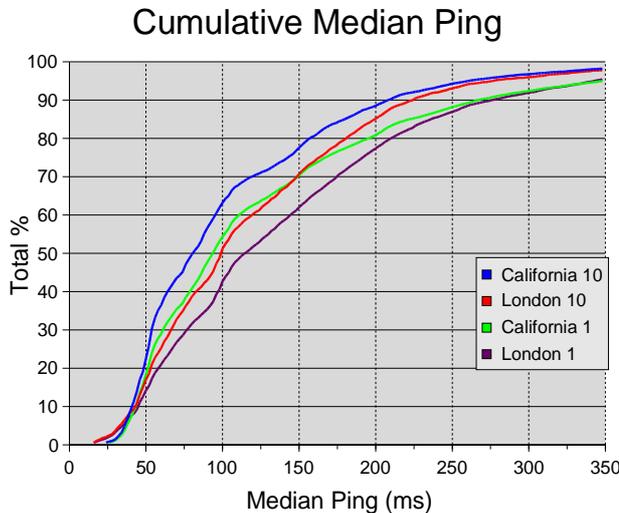
Figure 4 Cumulative median ping per player per game

A game-hosting service provider should be most interested in the tolerance of their most demanding customers, so it is reasonable to pay closer attention to the rate at which active players come back to servers. For each server there are two plots - one marked "10" where we excluded players who picked up less than 10 items per minute, and the other marked "1" which includes all players who picked up at least 1 item per minute. (The number of items picked up per minute is taken to be indicative of how actively the player is actually involved in the game. As noted above, players who accumulated less than 10 ping samples in a game were not counted for that game.)

The California and London curves are fairly similar. Considering the minimal activity (at least one item per minute) curves first, we see that 80% of the player-games on the Californian server were played by clients with a ping of less than 196 milliseconds. The equivalent threshold on the London server was approximately 210 milliseconds.

When we consider only players who picked up at least 10 items per minute (reasonably active) the 80% thresholds drop to 158 milliseconds (Calfornian server) and 182 milleseconds (London server). If we take a slightly more aggressive stance and looked at the latency below which 70% of the player-games occurred, the tolerance would drop further to roughly 120 and 148 milliseconds respectively.

V. ANALYSIS

It might be argued that our results only prove that the Internet itself is bounded at a radius of something like 300-350 milliseconds. Therefore it is important to see evidence of topological locality in each server's player

populations before drawing specific conclusions from Figure 4.

A. Evidence of Topological Locality

Figure 5 and Figure 6 provide one source of evidence for regional locality in the player populations. The popularity of each server fluctuates on a daily and weekly basis in line with the afternoon and evening periods of the people who are geographically close to each server. In other words, given that London is 8 hours ahead of Palo Alto we see that the cyclical usage patterns are very similar and yet suggest rather different and distinct player populations.
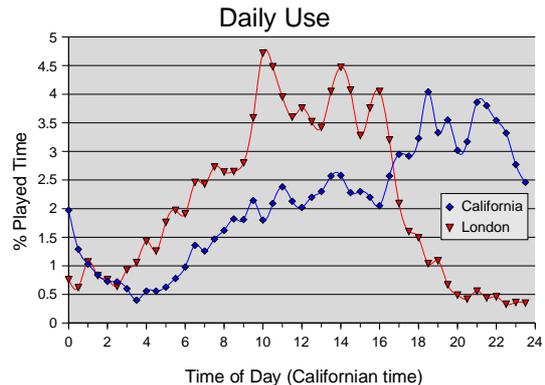
Figure 5 Daily playing time cycle for each server

(Note that in reality most of the Internet's topology is broadly based along geographical constraints. Thus these two figures merely confirm our expectation that players near London's timezone would see the London server as far closer to their latency tolerance than the Calfornian server. Conversely players in the US West Coast or Mid-West region would find the Calfornian server more attractive.)
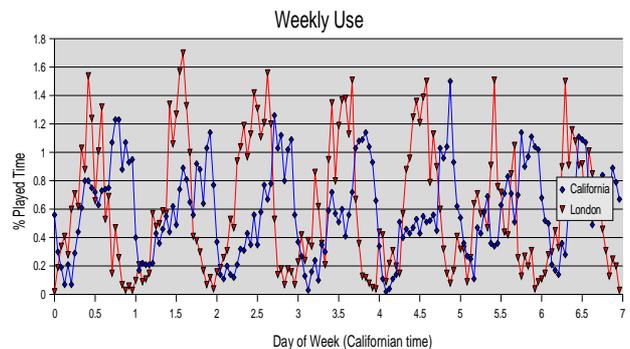
Figure 6 Weekly playing time cycle for each server

The IP addresses of each player provides an additional source of information about the player demographics. Reverse lookups of client IP addresses show that the Calfornian server was mostly frequented by North American ISP customers, while the London server saw European and US East Coast ISP customers. (The top domains are shown in Figure 7.)

We think there is enough evidence of regional and topological locality in the player populations of each server to justify treating the curves in Figure 4 as indicative of self-selection based on a server's latency to the client.

Rank	Calfirinia Games/Time (min)	Calfirinia Origin	London Games/Time	London Origin
1	323 / 3005	.ed.shawcable.net	108 / 1027	.pit.adelphia.net
2	192 / 2072	.cruzio.com	73 / 690	.Uni-Mainz.DE
3	124 / 1383	(RogersEAST/@Home)	75 / 679	.upc-d.chello.nl
4	119 / 1246	.018.popsite.net	50 / 606	(telnordia.se)
5	118 / 1221	.tx.home.com	53 / 604	.dyn.optonline.net
6	150 / 1200	.mediaone.net	44 / 565	(Rogers EAST/@Home)
7	132 / 1178	.pit.adelphia.net	35 / 463	.dyn.optonline.net
8	115 / 1151	.socal.rr.com	53 / 448	.dialup.tiscalinet.it
9	87 / 980	.pa.home.com	34 / 430	.pa.home.com
10	93 / 938	.sfba.home.com	20 / 288	.tx.home.com
11	69 / 799	.hsia.telus.net	24 / 273	.btinternet.com

Figure 7 Top 10 domains on each server

B. Why is this important

It is easy to qualitatively assert that latency sensitivity is important for players, and therfore is important for service providers who wish to understand how best to position their servers and market possible 'for-fee' game services in the future. The contribution of this paper is to provide some quantitative bounds to such a discussion. Armed with a network radius of around 150-180 milliseconds a service provider can place their game servers at optimal locations on the Internet relative to their target player market. Alternatively, if server placement is constrained by other business considerations, a 'radius' in milliseconds can help a service provider more accurately identify their likely customers and target advertising and support appropriately.
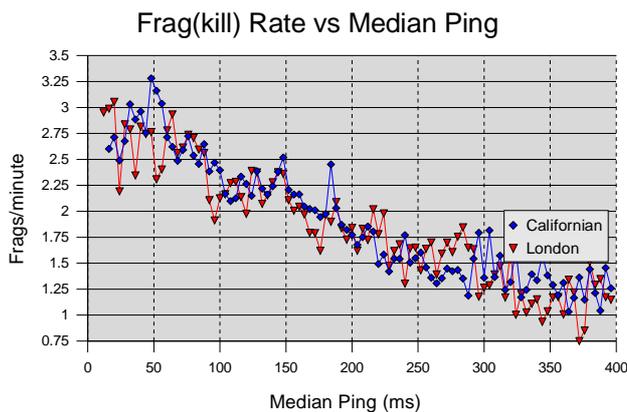
Figure 8 Frags per second as a function of latency

Our server logs also allow us to see the importance of latency in a quite dramatic manner. Qualitatively it is obvious that players with lower latency fare better. Figure 8 suggests more quantitatively that players with 45 milliseconds median ping were averaging 1 frag (kill) per minute more than a player with 200 millisecond

median ping. Given games running tens of minutes, this represents a not-insignificant impact on a player's game experience.

C. Limitations of the research

This analysis needs further development in a number of areas. Firstly, there is no data to indicate how important network jitter is to the game playing experience. It could be argued that higher latencies are associated with longer network paths between clients and server, and thus also equate to paths with greater jitter. Further work is needed to evaluate how sensitive fast, interactive online games are to jitter as distinct from absolute latency. A similar case can be made for investigating the sensitivity of our results to packet loss rates, which were also not measured.

VI. CONCLUSIONS

This paper provides a preliminary estimate of latency sensitivity amongst players of highly interactive, online multiplayer games such as Quake 3. Such quantitative estimates will become more valuable as 'for fee' game service providers try to understand where their game-playing customer base is located, or where best to locate their game servers for maximum customer satisfaction.

Two Quake 3 servers were instrumented to track their players and player's ping times over a period of months, and the results used to estimate a typical player's tolerance to network latency. The servers were placed 147 milliseconds apart and in different timezones - Calfornia and London. Both servers were configured identically, and advertised themselves to the Quake 3 online community as being both in Palo Alto, Calfornia. The only distinguishing feature from a player's perspective would be the ping time to each server.

The Californian server ran from May 17 to August 18, 2001 and saw 5290 unique clients who accumulated a total of 164 'days played'. 338 clients accumulated more than 2 hours total playing time during this period. The London server ran from May 29 to September 12, 2001 and saw 4232 unique clients who accumulated a total of 77 'days played'. 131 clients accumulated more than 2 hours total playing time during this period.

The player populations of each server were seen to be distinct, based on playing times being similarly cyclical on a daily and weekly basis yet phase shifted 8 hours relative to each other (the timezone difference between California and London). Analysis of the client IP addresses confirmed each server appealing to clients whose ISPs were closer topologically and geographically.

We plotted the cumulative distribution of median-ping per player per game as an indication of how frequently each server was visited as a function of experienced latency. Given that the player populations appear to have conciously self-selected based on ping times, we feel comfortable that our results suggest players prefer servers less than 150 to 180 milliseconds away. Other things being equal, players will migrate away from servers outside this range.

Additional work is required to further understand and quantify the impact of network jitter to a player's online experience.

REFERENCES

[1]   "Quake 3", http://www.idsoftware.com/

[2]   "Half-Life", http://www.valvesoftware.com/projects.htm

[3]   Henderson, T., Latency and user behaviour on a multiplayer games server. Proceedings of NGC 2001, London, UK, pp1-13, November 2001

[4]   Sheldon, N., E. Girard, S. Borg, M. Claypool, E. Agu, "The Effect of Latency on User Performance in Warcraft III," Technical Report WPI-CS-TR-03-07, Computer Science Department, Worcester Polytechnic Institute, March 2003

[5]   Henderson, T., S. Bhatti, "Modelling user behaviour in networked games," Proceedings of ACM Multimedia 2001, Ottawa, Canada, pp212-220, October 2001

[6]   Wu-chang Feng, Francis Chang, Wu-chi Feng, Jonathan Walpole., "Provisioning On-line Games: A Traffic Analysis of a Busy Counter-Strike Server," SIGCOMM Internet Measurement Workshop, November 2002

[7]   Johannes Färber, "Network game traffic modelling," Proceedings of the first ACM workshop on Network and system support for games April 2002

[8]   M. S. Borella. "Source models of network game traffic," Proceedings of networld+interop '99, Las Vegas, NV, May 1999

[9]   Gamespy3D, http://www.gamespy3d.com/

[10]   Armitage, G.J., "Sensitivity of Quake III Players to Network Latency," SIGCOMM Internet Measurement Workshop (poster presentation), San Francisco, November 1, 2001