Critical Sections in Networked Games

Saptarshi Debroy[1], Mohammad Zubair Ahmad[1], Mukundan Iyengar[2] and Mainak Chatterjee[1]
[1] Department of EECS, University of Central Florida, Orlando, FL 32816
Emails: {saptarsh, zubair, mainak}@eecs.ucf.edu
[2] Department of ECE, Stevens Institute of Technology, Hoboken, NJ 07030
Email: miyengar@stevens.edu

***Abstract*—This work introduces the concept of *critical sections* for online first person shooter games (FPS). A critical section is a section of game-play which demands higher precision or tighter deadlines. Critical section traffic is more sensitive to network degradations than sections immediately preceding or following it. Critical sections provide game developers and network programmers a notion of relative priority of game traffic, and can identify segments whose preservation can lead to superior user perceived quality of playing FPS games on a network.**

By analyzing video-recordings of over 5 hours of FPS game-play by 10 volunteers, we identify sections of FPS game-play whose degradation would cause inconsistent game-state updates resulting in user frustration. We observe that critical sections exhibit a pattern of occurrence and can account for upto 17% of game-play time.

We next quantify the expected network induced degradations on critical sections for online FPS games on the Internet. Using traces from a deployment of FPS workloads on 50+ nodes in the Internet, we study network dynamics and their ensuing effect on critical sections. Using traces from this experiment, we derive the lower bound on potentially degraded game-play session on todays Internet. We argue that critical sections of FPS games can be preserved. This can allow a variety of network architectures to better deliver higher perceptual experience when deployed on the Internet. Overall, our results have implications for FPS game-design, network provisioning, and game quality evaluation.

I. Introduction

"*I shot at him 10 times yet he killed me in just one hit*" – summarizes the frustrations of many online gamers when network degradations hamper game-changing situations. The success of a mission in a first person shooter (FPS) game often revolves around a few *segments* which are more critical to the outcome of the game compared to others.

FPS games are arguably one of the most popular game genres, with the latest edition of Call of Duty, Black Ops II reaching $1 billion in sales within 15 days of release [11]. In FPS, clients view and interact with the game world through an avatar in a first person setting. Typical online FPS games have missions with well defined objectives. Missions involve a series of time critical interactions which demand accurate game state information. Multiplayer games on the Internet primarily follow the client-server model of communication: there is a server which listens on a well known port for incoming client requests. These servers could either be run by the game company themselves, third party vendors who 'host' these dedicated servers, or by clients themselves who act as servers allowing them to invite friends and other game players.

Online FPS games have stringent service level agreement (SLA) requirements. FPS games are highly sensitive to delay, loss and jitter in the network. Many existing architectures on the Internet (e.g., DiffServ, Overlay networks etc) can help meet SLA requirements of delay and loss. However, such mechanisms cannot control packet loss or re-ordering due to lower layer connectivity issues, and can result in impaired game-state updates and user frustration.

A. Related Work

Prior work in this domain has involved client and server side measurements to characterize game traffic [4], [8]. Game traffic tends to exhibit a predictability, primarily consisting of small, periodic updates. The *effects* of network induced degradation on perceived game playing experience has also been investigated [3], [9]. Latency, loss, and jitter degradations on game playing experience have been evaluated by inducing these degradations in a controlled environment and asking users or computer agents to play in such conditions [3], [10]. The effects of such network induced degradations on subjective perception of a game player is studied in [6]. The correlation of player departure to network QoS levels was investigated in [5], where the authors demonstrate a strong correlation between player departure and inferred network degradation. However, much of prior work has treated *all* game segments to be of equal importance when evaluating perceptual quality.

B. Contribution

This paper introduces the concept of *critical sections* in online FPS games, and quantifies its occurrence and importance for subjective quality evaluation. We formally define critical sections to be sections of a game which contain events demanding higher precision and/or tighter deadlines. To identify critical sections, we recorded multiplayer game-play of 10 volunteers, which was played back and rated to identify sections of the game which we think were critical. Recordings of several matches in two popular games of *Call of Duty* (CoD) series played by our subjects for a total of 5+ hours serves as an input to our study. We observe that critical section occurrences are bursty in nature, exhibit a certain pattern of occurrence and can account for anywhere between 11% to 17% of total game-play time.

We then analyze network level degradations of todays Internet that impacts critical sections. As an input to our study, we collect traces from a deployment of dummy clients and servers

978-1-4673-3122-7/13/$31.00 ©2013 IEEE

(a) (b)

Fig. 1. Examples of critical sections: (a) A sniper shot requiring high precision, (b) Jumping from a chopper within a tight deadline

on the Internet using 50+ PlanetLab vantage points with a presence in the US, Europe, and Asia-Pacific. Client-server paths are probed according to real game traffic patterns. Using this study, we quantify the vulnerability of critical sections to network induced degradations. We observe that a FPS game-play section is critical with probability 0.17, and that at least 10–20% of the time a player can expect to have degraded game-play session. Overall, our results have implications for FPS game-design, network provisioning, and game quality evaluation.

The rest of the paper is organized as followed. Section II defines and quantifies *critical section*. In Section III, we present the our network measurement methodology. We discuss the limitations of our experiments and measurements in Section IV. We draw conclusions of our discussion in Section V.

II. Critical Section

QoS has long been used to measure network goodness for interactive multimedia applications. However, QoS lacks an important aspect in quantifying performance, largely due to the void between system centric and human centric performance evaluation. We argue that quantification of QoS alone says little about game playing *experience*. This is because not all phases of game playing are equally critical.

A certain amount of network degradation can result in completely different levels of subjective perception depending on the sensitivity of player action during game-play. Game-play can include actions like using a sniper, lobbing a grenade, crouching or moving within a map, consuming health potions and so on. As the authors in [7] point out, player actions with tighter *deadlines* demanding high precision are clearly more important than other game-play sections such as exploring a region or inspecting elements. We correspond to such segments as *critical sections*. Events in *critical sections* are more sensitive to network level degradations and have more control on the game outcome than sections immediately preceding or following it.

A. What is critical?

The concept of *critical section* is highly subjective and dependent on player actions. In a typical FPS game, player actions broadly consist of movements, engaging in combat

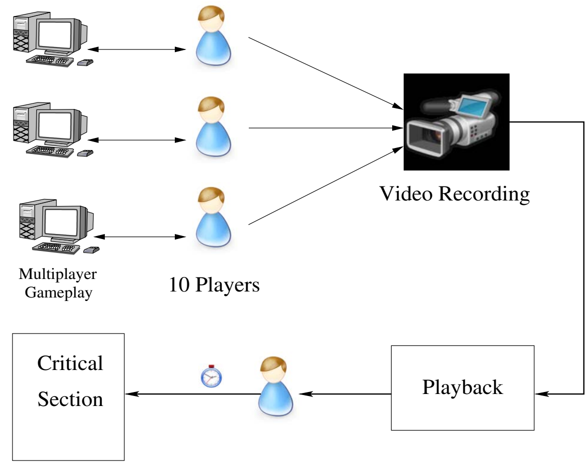


Fig. 2. Experimental workflow for recording critical sections

and performing special tasks. Intuitively, any game session involving combat is more sensitive to network degradations than a session which only involves movements. However, if a tighter deadline or a higher precision is associated to movements, its 'degree of criticality' increases. For example, a game session which requires a jump from/onto a chopper associated with a tight deadline is a *critical section* (Fig. 1(b)). We further argue that not all sections of a *critical section* are of same different degree of criticality; it varies depending upon the accuracy of game state information that the section requires. For example, sniping the enemy requires a much precise knowledge of the target than 'lobbing a grenade' to the enemy bunker; shooting with a 'M14 EBR' demands more accuracy than shooting with a 'AK-47'.

We make a distinction between critical sections with tighter deadlines and sections which require critical *tactical* decision making. For example, consider a scenario when a player goes around a building to reach a rendezvous point, only to realize that the path she takes often leads to the player missing the deadline. As the player repeats the mission, she realizes that the only work around to this is when she takes an alternative (shorter/non-obvious) route to the rendezvous point. In this case, the player missing the deadline is a result of a *tactical* decision which is not a true critical section. In other words, network degradations that can occur during a tactical decision making phase are clearly not as critical as a combat scenario or a sniper shooting event.

B. Experimental Setup

Survey Subjects: We asked a group of 10 volunteers to play and record a variety of missions of various FPS games. The subjects chosen were all male and generally in the age group of 20-30 years. These subjects responded to an advertisement calling for gamers, and were versed with playing online FPS games.

Experiment Workflow: Our set-up consists of a series of networked machines which host a variety of FPS games with different multiplayer missions within them. On an average, each player accounts for about 30 minutes of game-play, which

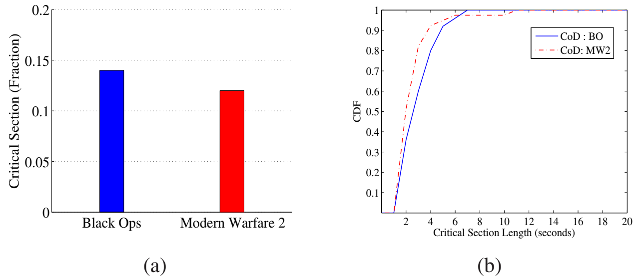


Fig. 3. (a) Fraction of time spent in critical section, (b) CDF of critical section length

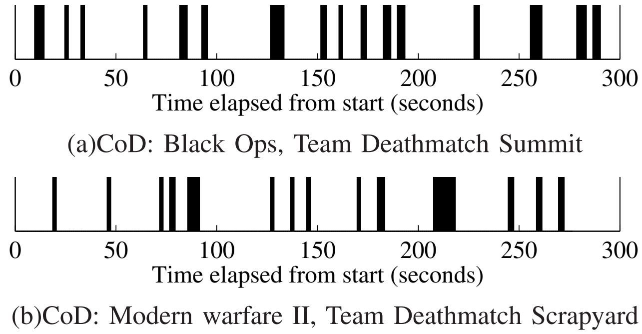


Fig. 4. Distribution of critical sections in game missions

amounts to over 5 hours of collective play. Fig. 2 provides a high-level overview of the experiment workflow. We provided the gamers with a choice of two popular Call of Duty (CoD) games: *Modern Warfare II* and *Black Ops*. The choice of the games reflects variety in war-scenario, weapons-of-choice, strategies, and terrains. These range from World War II (old-fashioned weapons), the Cold War to Modern Warfare (more evolved weapons). Further, carefully selected matches within each game challenges players in different game objectives (demolition to protection, escape to invasion) and hardcore 'deathmatches'.

While some players had existing foreknowledge about certain missions, others did not. To eliminate bias, we asked the game-players to familiarize themselves with a given game before recording game-play. Game sessions were recorded on the terminal for offline analysis.

Post-Analysis: The recordings of these games were then analyzed and critical section start/end time was recorded with a stopwatch during playback.

C. Results

We begin with some high level results. Next we discuss the distribution of *critical section* duration for different games. We also show frequency of *critical section* bursts experienced by gamers playing different missions in the three games. We round off the discussion with some emerging trends in playing styles and strategies used.

High Level Result: Fig. 3(a) shows high level results from this round of experiment. The plot records the fraction of total game-play time that involved critical section which either had high precision requirements or tight deadlines, varying from 0.11 to 0.17 for the different matches. The two extremes correspond to a 'capture the flag' operation and a 'demolition' operation respectively. The plot shows that roughly 80% of game traffic which is non-critical can be deemed *insensitive* to network induced degradations. In other words, even if game traffic carrying this portion of (non-critical) game-state is corrupt, it would lead to a fairly consistent user perceived quality.

Critical Section Duration: Fig. 3(b) shows the distribution of critical section duration for different matches. We find that more than 90% of all the critical sections last between 1–4 seconds for CoD: Modern Warfare II. Critical section bursts of same length comprise about 70% of all critical sections for CoD: Black Ops. Bursts lasting more than 6 seconds correspond to Airstrike, Predator Drone attacks etc. which increase player's killstreak.

Recurring Frequency: To further analyze the frequency of critical section occurrence, we present in Fig. 4 the binary presence of critical sections throughout a mission. The time distributions reveal the randomness of the *bursts* which signifies that critical section occurrence during a match is dependent on player actions. However, a trend of small durations of critical sections separated by larger bursts of non-critical sections is evident.

Game Tempo: In Fig. 5, we show expected tempo of playing a mission in a typical FPS in terms of probability of *critical section* occurrence. It shows that occurrence of *critical sections* is not equally likely throughout the mission. Higher peaks at the end suggest that a gamer can expect more frequent bursts of *critical sections* at later stages of a mission. Also there are always quieter non-critical sections at start and end of a mission mostly comprising of cut-scenes, auto play modes, mission briefing etc.

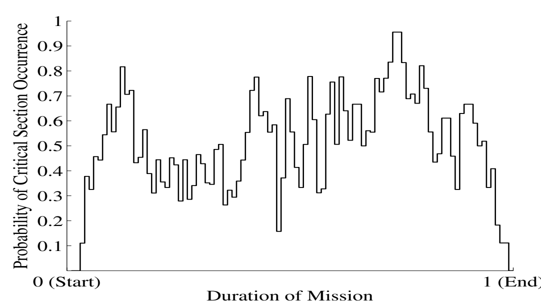


Fig. 5. Probability of critical section occurrence in a FPS

Post analysis and player interviews reveal that players playing against other players very seldom tend to get engaged in prolonged gun-battles. Critical sections with very small duration (≈1 second) represent 'kills' occurring due to blind-sidedness of the opponent resulting in precise headshots. Most of the other critical sections (duration of 2–4 seconds) are due to bullet exchanges with sub-machine guns (SMGs) requiring longer period due to their low accuracy. However, such SMGs are popular for invasion purposes for high rate of shots. Players also tend to use high accuracy, low rate LMGs (Low Machine Gun) while defending strong tactical positions resulting thinner bursts.

D. Implications

Our results have implications for game-designers with respect to optimizing traffic during critical sections, as well as tolerating slack when such sessions are in progress.

Packet Tagging: Tagging critical section packets can help network elements identify traffic whose delivery is relatively more important. Packet tagging can be done in the application layer while executing the game code itself or using a real time statistical prediction model. While running the game code, execution of event handlers (piece of code representing gun-shot, lobbing grenade etc.) responsible for critical section can initiate tagging of packets for preferential treatment (differential QoS), or transmitting redundant copies on multiple paths. Since this amounts to less than 20% of total game-play, we believe that it is sufficient to perform network optimizations for only one fifth of total game traffic. This translates to a cost reduction of more than 80% in preferentially servicing critical section FPS game traffic compared to the entire game traffic.

Game Slack: Our results also have implications for game designers. If game design tolerates little or no slack in servicing game traffic during critical sections, simple network degradations can often result in extreme user frustrations. Players unanimously agree to the fact that they play to win: degradations that lead to a player losing in a mission are deemed extremely frustrating. Game designers who aim to provide increased user satisfaction could additionally tolerate a higher threshold of slack during critical sections.

Quality Evaluation: Identification of critical sections can help perform accurate subjective quality evaluation. Quality evaluation of FPS games often involves playing games on a LAN, where losses can be induced in a controlled manner. This mode of evaluation treats every game-segment alike. We argue that quality evaluation is best performed by inducing degradations on critical sections rather than the entire game traffic.

III. Network Measurement

We round off our study by answering the following question: how vulnerable are critical sections in todays Internet?

Methodology: To emulate the prevalent client-server architecture in typical gaming scenarios over the Internet, we carry out a set of probe based measurements using geographically diverse PlanetLab [12] vantage points. Fig. 6 provides a high-level overview of a typical measurement architecture. We designate a set of 16 nodes as servers and 40 nodes as clients located all across the globe. The servers send probe packets (pings in our case) to each client and vice versa over a period of 10 minutes (typical length of a multiplayer 'deathmatch') for every 30 minutes. Thus there are two probing sessions per hour which is carried out for an entire day to give us 48 gaming sessions. During every session, the size of the probes are varied according to the distributions mentioned in [1], [2]. This characteristic enables us to roughly mimic traffic generation patterns in well known FPS games. While ping probes do not provide accurate values of latencies/losses along an Internet path, it provides us with a decent estimate of the path characteristics and we aim to further refine our experiments with actual gaming traffic in future work.

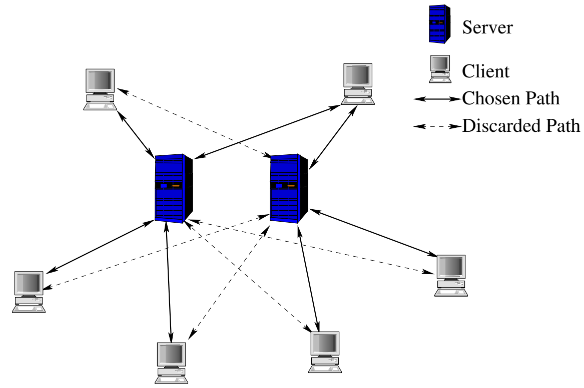
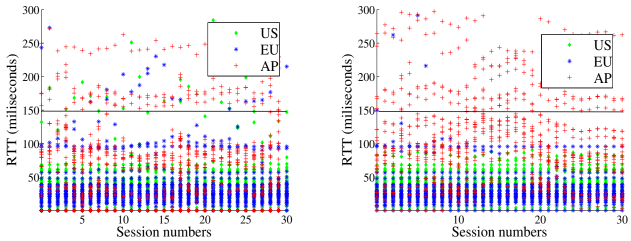


Fig. 6. Sample experimental scenario with 2 servers and 6 clients

We measure total path degradation obtained for each gaming session across both client-server and server-client paths. Our criteria for declaring a path degradation is an RTT exceeding 150ms in accordance with researchers in [3], [7], [10]. We do not consider server-client paths with default latencies greater than 300ms over all sessions (these paths are indicative of far away locations between server-clients). Figure 6 presents an example experimental scenario where we have 2 servers and 6 clients. The discarded paths represent those paths which are not considered since their default average end-to-end latencies are greater than the accepted maximum in the case of gaming sessions (greater than 150ms).

Quantifying degradation: Increased interactive delay for FPS games can *severely* impact *critical sections*, and is arguably the most important QoS parameter to be considered. This is because a rise in delay corresponds to decreased interactivity. This includes slow game updates for actions (from server to client) and inputs by players (from client to server) causing inconsistent game play updates, and eventual subscriber churn.

Basic latency measurements from server to client and client to server paths are presented in Fig. 7(a) and Fig. 7(b) respectively. We take the average RTTs for sessions of length 1 minute which indicates the maximum duration of a typical *critical section* burst. We infer RTT values of above 150 ms as degrading, and often notice spikes in various segments of critical sections. These spikes do not indicate instances but entire sessions of degradations potentially resulting in very poor perceptual experience. In Fig. 7(a), paths of all the continents show instances of degradation. However in

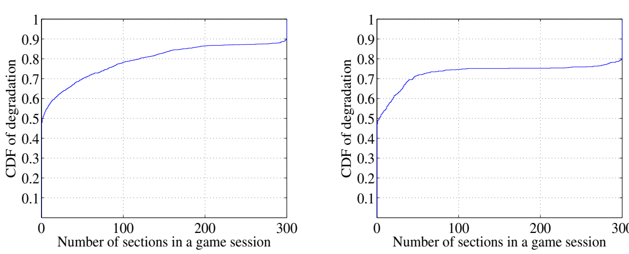


(a) (b)

Fig. 7. (a) Server to client average RTT distributions, (b) Client to server average RTT distributions

Fig. 7(b), we see the lion share of degraded paths from AP with US being fewest. Also the two figures separately give a lower bound of total instances of degraded sessions; to get the actual number of degraded session, we need to combine both. For example, the average RTT from AP server to AP client number 14 for session number 10 is below the threshold of 150 ms. But the reverse path show average RTT above 150 ms resulting session number 10 to be degraded.

Probabilities of possible game session degradations from server-client and client-server paths are presented in Fig. 8(a) and Fig. 8(b) respectively. We divided each 10 minutes long game play sessions into 300 sections of 2 seconds duration as it typifies more than 50% of critical section length. Such divisions are done for all paths. Figure 8(a) shows that for server to client paths about 10% of all game-play sessions are expected to have degraded critical sections. The same metric for client to server paths is around 20% of all game play sessions shown in Fig. 8(b). This means that on todays Internet an online game player can expect a degraded game-play session with a lower bound probability of 0.1 to 0.2.



(a) (b)

Fig. 8. (a) Degradations for server to client paths, (b) Degradations for client to server paths

IV. Limitations & Future Work

Our experimental workflow and resulting conclusions are not devoid of certain limitations. We discuss the limitations and corresponding future works.

Type of FPS: Our evaluation is limited to a certain type of FPS games (CoD series). Though CoD is a popular FPS game that deals with war-like scenarios, we plan to study stealth and team combat FPS games like CounterStrike, Unreal Tournament, Quake etc. in our future surveys.

Subjective Perception: We believe critical sections can act as a useful aid in subjectively assessing game-play quality. In a future study, we would additionally like to quantify the effects of network induced degradation on critical sections to truly ground our ideas.

V. Conclusions

This paper presented and evaluated the concept of critical sections in online FPS games. A critical section is a section of game play which is more sensitive to network degradations than sections immediately preceding or following it. Critical sections can provide network administrators and game developers a notion of relative priority within a game, allowing for a host of optimizations to improve game-playing quality. To quantify critical sections, we recorded game-play of 10 volunteers playing different multiplayer matches in two different Call of Duty games amounting to over 5 hours of game-play recordings. Using this as an input, we characterized offline the distribution, recurring frequency, and percentage of game play that is critical. Using real world Internet traces, we show that critical section traffic gets degraded when played over Internet. The goal of this work is to move towards a generic framework that can automatically tag packets that carry critical section content, making them visible to various layers for optimized game performance. Our characterization of critical section can serve as an useful input for network protocol optimization, game design, as well as game-quality evaluation.

References

[1] G. Armitage, "At The Intersection Of Networks And Highly Interactive Online Games", in Algorithms for Next Generation Networks 1st edition, G. Cormode, M. Thottan (Eds.), Springer, 2010, pp. 403–434

[2] G.Armitage, M.Claypool, and P.Branch, "Networking and Online Games- Understanding and Engineering Multiplayer Internet Games", John Wiley & Sons, UK, April 2006

[3] T. Beigbeder, R. Coughlan, C. Lusher, J. Plunkett, E. Agu, and M. Claypool, "The effects of loss and latency on user performance in unreal tournament 2003", *Proc. ACM SIGCOMM NetGames*, 2004.

[4] P.A. Branch, A.L Cricenti, and G.J. Armitage, "A Markov Model of Server to Client Ip traffic in First person Shooter games", *Proc. IEEE ICC*, May, 2008.

[5] K. T. Chen, P. Huang, G. S. Wang, C. Y. Huang, and C. L. Lei, "On the Sensitivity of Online Game Playing Time to Network QoS", *Proc. IEEE Infocom*, March, 2006.

[6] P. Chen, M. El Zarki, "Perceptual view inconsistency: An objective evaluation framework for online game quality of experience (QoE)", *Proc. ACM NetGames*, 2011.

[7] M. Claypool and K. Claypool, "Latency can Kill: Precision and Deadline in Online Games", *Proc. ACM Multimedia Systems Conference (MMSys)*, Scottsdale, AZ. Feb. 2010.

[8] W-C. Feng, F. Chang, W-C Feng, and J. Walpole, "A Traffic Characterization of Popular Online games", *IEEE/ACM Transaction on Networking*, 2005.

[9] T. Henderson and S. Bhatti, "Modelling user behavior in networked games", *Proc. ACM Multimedia*, Oct. 2001.

[10] T. Henderson and S. Bhatti, "Networked games: a QoS-sensitive application for QoS-insensitive users?", *Proc. SIGCOMM RIPQoS*, 2003.

[11] Call of Duty: Black Ops II. www.callofduty.com/blackops2

[12] PlanetLab Consortium, www.planet-lab.org