A Measurement Study of Virtual Populations in Massively Multiplayer Online Games

Daniel Pittman and Chris GauthierDickey
Department of Computer Science
University of Denver
{dpittman | chrisg}@cs.du.edu

ABSTRACT

Understanding the distributions and behaviors of players within Massively Multiplayer Online Games (MMOGs) is essential for research in scalable architectures for these systems. We provide the first look into this problem through a measurement study on one of the most popular MMOGs, World of Warcraft [15]. Our goal is to answer four fundamental questions: how does the population of the virtual world change over time, how are players distributed in the virtual world, how much churn occurs with players, and how do they move in the virtual world. Through probing-based measurements, our preliminary results show that populations fluctuate according to a *prime-time* schedule, player distribution and churn appears to occur on a power-law distribution, and players move to only a small number of zones during each playing session. The ultimate goal of our research is to design an accurate player model for MMOGs so that future research can predict and simulate player behavior and population fluctuations over time.

1. INTRODUCTION

Since the introduction of the first massively multiplayer online games (MMOGs) in the 1990s, MMOGs have continued to gather larger audiences and generate more revenue for companies each year. For instance, World of Warcraft, currently one of the most popular MMOGs, has over 9 million subscribers and generates over $1 billion in gross income each year. With their increased popularity has come an increased interest in the research community on the design of scalable architectures for these systems. While most research has focused on distributed and peer-to-peer architectures for scalability purposes, data from real MMOGs beyond network traffic has not been measured and realistic models for simulations have not been used for validation.

The ultimate goal of our study is to design a realistic and empirically based simulation model from measurements taken from current, commercial MMOGs. This model could be used in the design and evaluation of future architectures for MMOGs, and furthermore the data we collect can be used as trace data to drive simulations. Previous research has measured network traffic characteristics, such as where players are located geographically and inter-packet arrival times [4, 9]. While understanding the network patterns exhibited by MMOGs is very important, we cannot use this data to help us model player movement or player distributions in the virtual world.

Instead, we measure the populations and behaviors of *virtual* populations entirely from within an MMOG. This allows us to not only collect data about the current census of the MMOG, but also allows us to monitor session times, what areas players visit, and how long they play in a given area.

We conjecture that at least four facets of the virtual populations and behaviors seen in MMOGs are required for a viable model. These four facets are: *1) Population changes over time:* both the magnitude of actual populations and the fluctuations from the minimum to the maximum census are needed to model the changes in populations over time; *2) Arrival rates and session duration of players:* the arrival rates and session lengths indicate how much *churn* occurs in the online player populations and can greatly affect the performance of a given architecture; *3) Spatial distribution of players over the virtual world:* as with arrival rates and session lengths, the spatial distribution of players can have a large impact on the performance of an architecture; *4) Movements of players over time:* how players move, such as how many distinct areas they visit and how long they remain in a given area affect underlying architectures. If players are in constant flux, then large-scale architectures will need to cope with constantly moving players.

We performed an extensive set of probing-based measurements over World of Warcraft which uses a client/server architecture typical of all commercial MMOGs. The client provides a Lua [12] scripting interface which allows end-users to create *add-ons* for the game. Using this interface, we scripted a set of measurement tools to record populations, distributions of players across zones, arrival and departure rates, session times, and player movements. In our measurements, we recorded over 5 weeks of data, over 12,000 individual players online and over 32,000 individual playing sessions.

Our preliminary results show that populations follow a diurnal schedule where peak playing time is during the evenings and weekends. We also obvserved that player distributions and session lengths in the virtual world appear to occur according to a power-law distribution. Finally, we see that players tend to visit only a small number of zones, and remain in each zone for a fairly significant portion of their playing time. *These results differ significantly from simulation parameters used in prior research that has analyzed MMOG architectures.* Thus, a major contribution of our work is a more realistic model for analyzing different MMOG architectures.

2. BACKGROUND

Scalable architectures for MMOGs have been an active area of research over the last several years [5, 1, 2, 10, 13, 8, 7]. These

NetGames'07, September 19–20, 2007, Melbourne, Australia

architectures are typically distributed, or peer-to-peer, because the client/server architecture requires a huge investment in resources to scale with the peak demand on a realm. For example, Sony Online Entertainment has approximately 150,000 players online at any time of the day, yet it requires approximately 1,500 servers, averaging only 100 players per server [11].

To validate these distributed architectures, researchers have used simulations with artificial workloads, with the exception of Baughman and Levine, who used real traces from a small networked multiplayer game called XPilot [1, 16]. The artificial workloads used by past researchers were similar and typically consisted of a virtual world evenly divided into zones with a uniform distribution of players across the virtual world. Player movement was simulated by having players remain in their zone for some uniformly distributed amount of time and then randomly choosing a new zone to travel to. *Our results show that these artificial workloads differ significantly from what we observed in our measurements.*

To date, little work has been done on the characteristics of player populations in MMOGs. Some websites have measured population sizes and other statistical game information from their favorite MMOG, including data such as the kind of equipment that players have, where treasures are located, and how often particular items re-spawn. These informal measurements are used by players for game strategy and are not part of a systematic and scientific measurement study. In particular, one add-on written for World of Warcraft, called Census+, was designed to capture game characteristics of the player database and record this information online. We modified this add-on and used it in conjunction with our own add-on to correctly measure virtual population information [3].

With respect to scientific measurement studies, Chambers et al. have studied various network conditions related to players and the server of small networked multiplayer games [4]. Similar to our data, their measurements also show diurnal patterns of game populations. Kim et al. measured network patterns on Lineage II, a popular MMOG in Korea [9]. Their work focused on network packet sizes, RTTs, session times and inter-session arrival times. Their data on session times appears to show a similar power-law distribution of playing times, with 50% of all playing sessions lasting less than 26 minutes. This data shows similar results to the session times we recorded. Tan et al. measured player mobility in small networked first-person shooter (FPS) games and designed a mobility model for that class of games which they call the Networked Game Mobility Model (NGMM) [14]. They showed that the typical random way-point mobility model is not sufficient for modeling player movements when compared to actual traces of player movements from FPS games.

Our study is the first systematic and scientific characterization of player populations and behaviors in commercial MMOGs. While prior work has focused on measuring network characteristics exhibited by multiplayer games, we instead directly measure user behavior in the virtual world so that we may characterize virtual populations of MMOGs.

3. MASSIVELY MULTIPLAYER ONLINE GAMES

Massively multiplayer online games (MMOGs) are a class of multiplayer computer games where several thousand players interact in the same virtual world. A typical MMOG might have several million subscribers, but these subscribers cannot all play in the same virtual world. Instead, an MMOG is divided into *realms*, which are copies of the same virtual world, each hosting several thousand registered players. The realms are then geographically distributed across the globe so that US players connect to US realms, European players connect to European realms, and so forth. As an analogy, one can think of the game Monopoly™. At any time, millions of people might be playing Monopoly, but each game can only support a maximum number of players. While each game has the same pieces and cards, the state of each game is different.

To date, all commercial MMOGs use a client/server architecture. Most commercial MMOGs are divided into *realms*. Each realm is managed by a logical server usually in the form of one or more physical machines acting as a cluster. Players in the MMOG play the game through an alter-ego, or character, represented on screen in a 3-dimensional world. They can interact with the world and with other players. A realm in an MMOG can handle several thousand players simultaneously. For example, in the popular MMOG called EverQuest [6], a typical cluster handles around 2,500 players concurrently, with around 10,000 players registered on each realm [11]. To compensate for millions of subscribers, companies host hundreds of realms across the world.

Realms are further divided into distinct areas called *zones*. Each zone ranges in difficulty, preventing new, less experienced players from traveling to more difficult zones. This subdivision helps prevent players from amassing in a single zone which would result in processing difficulties for the server. Indeed, the processing cost for the server for n nearby players could be $O(n^2)$ if it must consider interactions between all pairs of players that are within a close distance to each other.

4. MEASURING VIRTUAL POPULATIONS

Two methods can be used for measuring virtual populations and behaviors of players in MMOGs. The first method is to analyze logs generated directly from an MMOG. This method has the advantage of being accurate, though few companies are willing to share logs from their games and they may not contain the needed information. The second method is to use probing-based measurements to try to infer properties of the system. We use the second method in this study.

We measured data from World of Warcraft over 5 weeks from April 13th, 2007 to May 20th, 2007 for close to 24 hours a day using a measurement tool we designed. World of Warcraft is divided into close to 200 realms at the time of this writing. We measured data from a single realm (Aerie Peak), chosen randomly from a list of realms as having a full population[1]. We compared our measurements with a smaller set of measurements taken on a different randomly chosen realm and we found that the results were similar, indicating that our measurements are generalizable to other World of Warcraft realms. Once a day, our measurement tool logged out of the system so that data logs could be written to disk and archived, causing a small amount of measurements during this time (lasting approximately 10 minutes) to be lost.

In order to measure population information, we designed a set of scripts that run from the client using the Lua [12] scripting interface provided by World of Warcraft. We modified the Census+ add-on to more accurately record population snapshots and information to collect the data scientifically [3]. We also wrote an additional add-on to record continuous information about a randomly selected subset of players.

World of Warcraft provides an interface into the *who* service and provides a *friends* list, both which assist our measurements. The *who* service is a service that the game provides which lets us query

[1]Blizzard provides a web page that shows the current status of each realm, including populations in general terms at `http://www.worldofwarcraft.com/realmstatus`.

which players are currently online and returns additional information to us such as what zone they are in. For scalability reasons, the who service provided by the World of Warcraft servers only returns 50 results per query and only accepts 1 query every few seconds. However, the who service allows parameters which enable us to narrow our search results to specific subsets of players. As such, when we query and receive 50 results, we know that we need to subdivide our current query into a smaller set of players.

To query all players in the realm, we begin by asking for players of a specific level (for example, we ask for all players at level 70). If 50 results are returned, we repeat the query, but ask for players of a specific class (e.g., all players that are Warriors and level 70). Once again, if 50 results are returned, we then ask for all players of a specific level, class, and whose name starts with a particular letter (e.g., Warriors, level 70, with a name that starts with 'a'). This technique allows us to systematically search the entire set of players currently online.

We record the current population every 15 minutes by taking two back-to-back measurements from the who service. Due to the desire to not flood the server with constant *who* queries, we limit our measurements of the entire population to 15 minutes. In addition, World of Warcraft is divided into two *factions*, or groups of players. Belonging to one faction prevents you from querying about players in the other faction. Thus, to measure the entire population, we must measure data from both factions. As future work, we plan on extending our data collection to fully study both factions, various realms, and on different MMOGs. However, we feel these initial measurements are an important first step in understanding the characteristics of MMOG populations.

Because each measurement takes up to five minutes, the population is changing while we are measuring. The back-to-back measurements are designed to give us an idea of how much fluctuation is occurring while we are measuring. For example, the union of the two measurements represents the stable population over the measurement period. On the other hand, the players seen in the second snapshot that are not seen in the first snapshot represent the arrival rate over the time it takes to complete the second snapshot.

The second interface, the *friends* list, allows us to store updated information about 50 players in the game. World of Warcraft sends an event whenever one of our 'friends' logs into or out of the game. We can send a message to the server to update the current status of our friends list more frequently than we can measure the entire player database. Therefore, after a set of back-to-back measurements, we add up to 50 randomly selected new arrivals (those seen in the second snapshot that were not in the first snapshot) to our *friends* list. We then ask the server for a friends list update every 30 seconds. When a friend leaves, we replace them after the next back-to-back population measurement. This allows us to track the session lengths and movements of those players on our list with more accuracy. Thus, from the server we can record where a friend is every 30 seconds and exactly how long they remained in the game.

At the end of our measurement period, we measured how many unique players we saw to determine the accuracy of our sampling method. On the single server we measured, we saw over 12,000 unique players which accounts for almost half of the player database on a World of Warcraft realm, indicating our sampling method is reasonably accurate.

5. MEASUREMENTS

We divide our measurements into four sets: population sizes over time, arrival and departure rates, session lengths, and player distributions. Our first two sets of measurements indicate the volume of players and the amount of churn that occurs on a given MMOG server. The second two sets of measurements indicate how long players play and both how they move and are distributed in the virtual world. Together, these measurements can help us determine the behavior of virtual populations.

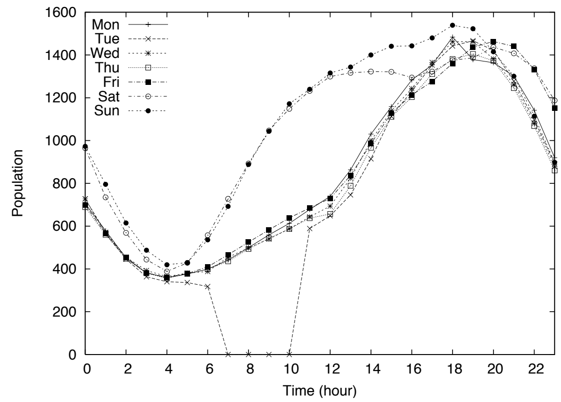

Figure 1: *Average daily population over time*: As expected, populations fluctuate according to a *prime-time* schedule, where evenings are more popular than days. From these measurements, we also see that players play earlier in the day on Saturdays and Sundays and stay up later on Fridays and Saturdays. Note that Tuesday from 6AM to 10AM is typically a *patch* day on World of Warcraft, where all servers are brought down for patching purposes.

5.1 Population Sizes over Time

In our first set of results, we examine measured population sizes over time, to understand fluctuations with respect to time of the day and day of the week. Our hypothesis was that more players were online during evening and weekend hours, due to weekly obligations such as work and school. Figure 1 shows weekly averages divided into days over the measurement period.

Figure 1 shows the 24 hour daily cycle with each line representing one of the days of the week. As expected, the server population increases significantly during the evenings and weekends. Note that the x-axis is in Mountain Standard Time, but players are not shown which time zone the realm is in when they choose a server to play on. Players in the United States typically play on US realms while European players play on European realms. Thus, we expect the populations to fluctuate with typical work and school schedules of the United States.

We note three important aspects of our graphs. First, populations peak at 1600 players. This occurs because this graph represent one faction and only levels from 1-69. From our measurements we have seen a 3:1 ratio between the two factions, indicating that if we had included both factions, the peak population would have been closer to 2100. In addition, we discovered that players who were level 70 exhibited different behavior from those who were levels 1-69. Thus, we excluded these players from the graphs. However, in future work, we plan on comparing the differences between the two sets of players. If we had included level 70 players, however, we would have seen close to 4000 players on an individual server. Second, we also see weekend play stands out significantly from weekday play. The realm experiences a significantly higher average population on Saturdays and Sundays, implying that the server must be provisioned to handle weekend play.

Last, we see an almost 5-fold increase in the number of players from the lowest point (at 4AM MST) to the highest point (18PM

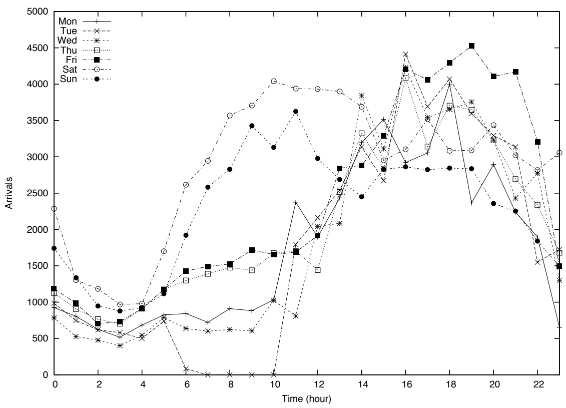

Figure 2: ***Arrival Rate*****: The arrival rate measures the number of new players we see from this hour that were not online during the previous hour and is directly related to the total realm population. As with the population graph, Saturdays and Sundays are the busiest days of the week. Note that during peak playing times, over 4000 new players log on each hour.**

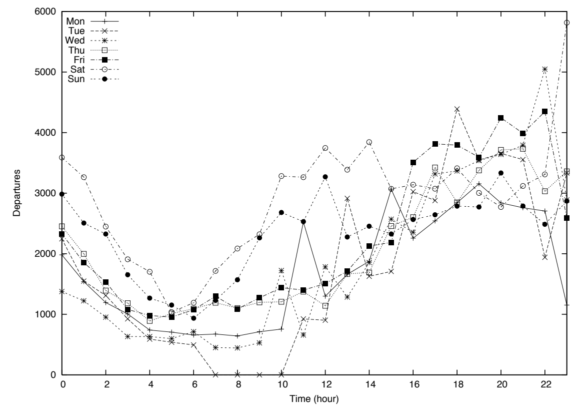

Figure 3: ***Departure Rate*****: The departure rate measures the number of players that were seen in the previous hour but are no longer online. Note that towards the end of the day, the departure rate increases dramatically (up to 6000 departures in an hour) as players are probably quitting the game to go to sleep.**

MST) of the realm population. This implies that servers must also be over-provisioned to handle peak loads during the evenings and are only partially loaded during the early mornings.

5.2 Arrival and Departure Rates

To further understand the population fluctuations and to help understand the amount of churn that occurs in a MMOG, we measured the number of arrivals and number of departures per hour and averaged this again on each day of the week. Figures 2 and 3 show these results.

In these two figures, we see that the amount of *churn*, or the number of players joining and leaving the game, is very high during peak playing times. Figure 2 shows similar trends of arrivals during the weekdays, but has an increased arrival rate on the weekends during earlier hours of the day. Figure 3 shows that the number of departures increases towards the end of the day. Together, *we see that during peak playing times, over 8000 players join and leave the game per hour.*

In terms of the magnitude of the difference between minimum and maximum arrival and departure rates, these results show that arrival rates differ by a factor of 4 while departure rates differ by a factor of 6. This implies that the distribution of players arriving during evening hours is more spread out. On the other hand, the results indicate that a large percentage of the population quits playing at the end of the night close to the same time.

5.3 Session Lengths

The third goal in our measurements is to understand the duration that people play the MMOG. Empirical evidence would suggest that players tend to log on and play for a long period of time. Anecdotal stories talk about players who play for days on end. However, our measurements in Figure 4 show that this is not entirely true.

Figure 4 shows the measured time playing versus the rank of the duration of play–that is, each session is ranked from the most number of minutes played, to the least. Using the *friends* list provided by World of Warcraft, we could track players to a granularity of approximately 30 seconds. From this figure, we see that only a small percentage of players we observed played for longer than 400 minutes (8 hours), while most players played for less than 200 minutes (3 hours). We did see, however, over the period of 5 weeks, that a handful of sessions were close to 24 hours. To understand this more clearly, we plot the CDF of the session duration in Figure 5.

Figure 5 shows that over 90% of the sessions were less than 200 minutes (or 3 hours) while 50% of the sessions were less 10 minutes or less. *This result verifies that MMOGs experience considerable churn. A large fraction of sessions are short lived while only a small fraction are stable.* We believe that what may be happening here is that players may be logging on to check to see if friends or guild members are currently online, checking in-game mail, or checking auctions at the auction house. If this is true, then the implication is that load on an architecture could be reduced by providing an external interface to these services that doesn't require logging into the game.

5.4 Distribution of Players

Our next set of measurements show the distribution of players in the virtual world. Recall that the world is statically divided into segments, called *zones*. We measure how many players are in each zone at 15 minute intervals. Figure 6 shows the number of players versus the rank of each zone from the most populated to the least and indicates a power-law relationship may exist. As the figure shows, a few zones have a large population above 40 players while a large number of zones have fewer than 10 players.

To further illustrate these measurements, we generated the CDF of the distribution of populations among zones. Figure 7 shows this distribution. These results indicate a non-uniform distribution of players as we predicted. Indeed, approximately 75% of the zones have fewer than 10 people in them, while about 5% of them have more than 40. Thus, the relationship between the number of players and the zones appears to be power-law. While this data does not take level 70 characters and the opposite faction into consideration, averaging those players into our data shows that the most popular zones have over 100 players in them.

5.5 Player Movement

The final aspect we measured was player movement. We want to investigate how many zones players visit during their playing time, and how long they remain in any given zone.

Figure 8 is the CDF of the distribution of players that will visit the number of zones indicated on the x-axis. These results show

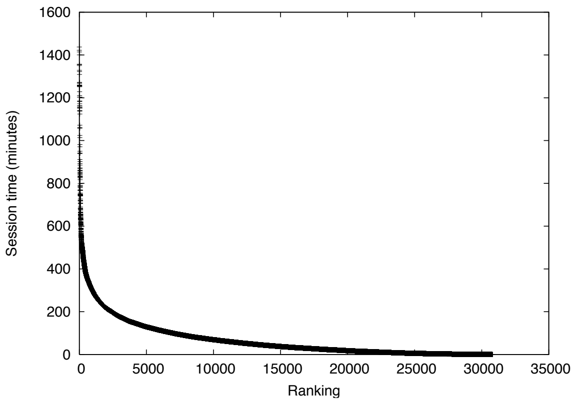

Figure 4: ***Session Times Observed*****: This measurement ranks all the sessions we measured (approximately 32k) and orders them from longest session observed to shortest session. This distribution appears to be power-law distributed; only a small fraction of players play for over 400 minutes, while the majority of players play for less than 200 minutes.**

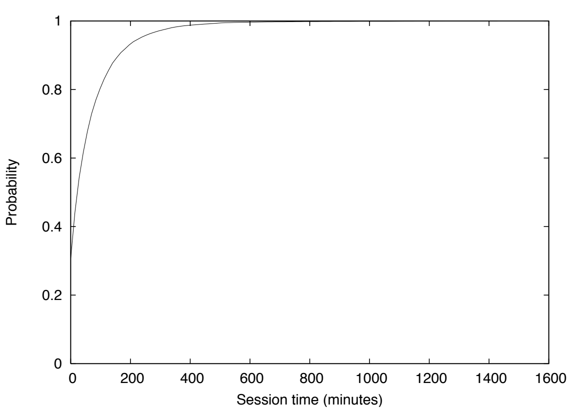

Figure 5: ***CDF of Session Times*****: This CDF clarifies the distribution in Figure 4. From this figure, we see that over 90% of the sessions were less than 200 minutes (3 hours). 50% of the sessions were less than 10 minutes.**

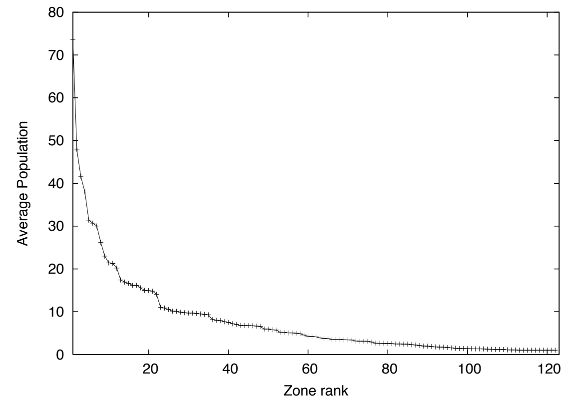

Figure 6: ***Distribution of players per zone*****: Zones are ranked from most populated to least populated to determine the distribution of players in the virtual world. From this result we see that a few zones have a large number of players while most zones have only a few players. Unlike prior simulation studies, this result shows that players are *not* uniformly distributed in the virtual world.**

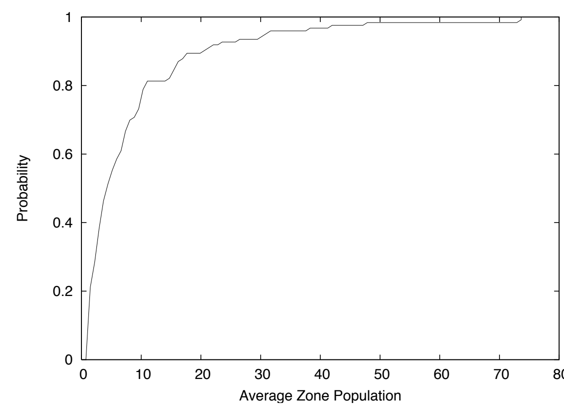

Figure 7: ***CDF of player distribution in zones*****: This CDF shows that approximately 75% of the zones have fewer than 10 people in them, while only a few have more than 40 people in them.**

that over 40% of the players only visited 1 zone while playing, while the majority of players visited 6 or fewer zones. This result is expected for two reasons. First, zones are tailored for players of a given level. Thus, players will typically only adventure in level-appropriate zones. Second, the previous results that shows half of the players only playing for 10 minutes or less. Thus, a player who logs on only briefly will probably only have time to visit one or two zones.

In addition, we measured how long a player remains in a zone during their playing session. This measurement allows us to see how mobile a player may be during their session, and hence how an architecture must handle player mobility. Figure 9 illustrates our results.

We see that similar to session times, half of the population spends less than 15 minutes in a zone. Because players visit only a small number of zones during their playing time, the average time they spend in a zone is naturally dependent on their session time. However, we can also see from these measurements that in general, players spend less than one hour in a zone.

6. CONCLUSION

Our initial measurements show clear trends in player population sizes, distributions and movement behaviors. First, like previous work that measured network traces in games [4, 9], populations fluctuate in diurnal patterns which appear to be correlated with typical prime-time media hours. The five-fold increase in population sizes from its lowest water-mark to its maximum indicates that server clusters must be over-provisioned to handle daily loads. If one is designing a client-server architecture, one should consider what uses the extra computational power during off-peak times could be used for. Unfortunately, delay constraints for interactive gaming may prevent one from using the extra server power for different time zones (i.e., such as helping with peak demands in an opposing time zone).

In terms of churn, we see from our arrival rates and departure rates that during peak playing time, approximately 8000 players joined and left the system in an hour. This is almost 2 to 3 times the total population seen on the server during that time, indicating that while some players may constitute a stable player base, a larger

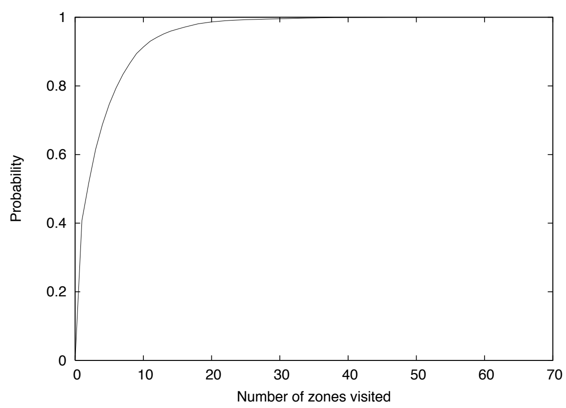

Figure 8: ***Number of zones visited*****: The percentage of players versus the number of zones visited demonstrates that most players only visit a small number of zones while playing. By examining the data, we see that most players visit 6 or fewer zones.**

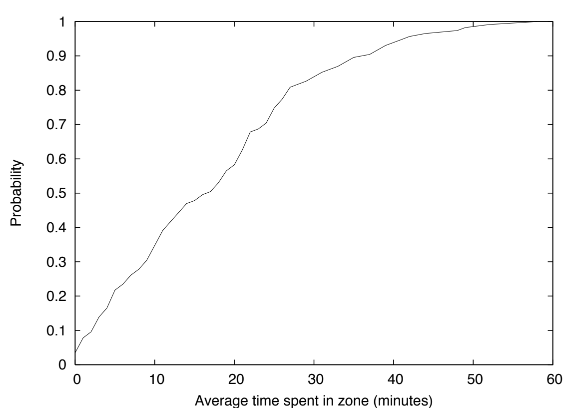

Figure 9: ***CDF of time per zone*****: In this graph, we measured the probability that a given amount of time would be spent in a zone. Our results show that over half the population spends less than 15 minutes in a zone.**

portion of the population is constantly being refreshed with new players.

Playing session durations appear to follow a power-law distribution where approximately 50% of the population remains online for 10 minutes or less. This impacts the design of MMOG architectures because it indicates frequent churn among players. Peer-to-peer architectures will need to ensure that the frequent churn does not result in a large amount of overhead for overlay maintenance. We hypothesize that players are connecting online to see if other friends or guild members are online and then leaving if not. We hope to test this hypothesis with future measurements.

The distribution of players in the virtual world also appears to follow a power-law distribution. This result is important because prior research has assumed the uniform distribution of players in the virtual world. We believe this result occurs from the design of MMOGs. In particular, certain locations which have rare equipment tend to be popular places to visit. Further, large cities which have a wide range of services also tend to be popular. Thus, players will tend to congregate in popular locations.

In summary, we see the following trends: 1) populations follow a sinusoidal curve with the minimal values in the early mornings and maximum values in the late evenings; 2) arrival and departure rates indicate significant churn during peak hours; 3) session times and player distributions in the virtual world appear to follow a power-law distribution; 4) the time spent per zone is dependent on session times, but is typically less than an hour.

As future work, we plan on measuring the geographic relationship among zones visited by a single player to understand what kind of mobility model we can use to simulate player movement. We also plan to incorporate all levels and both factions into our work. Finally, we plan to develop mathematical models to predict the distributions seen in our measurements so that player populations and behaviors can be accurately simulated.

7. REFERENCES

[1] N. E. Baughman and B. N. Levine. Cheat-proof playout for centralized and distributed online games. In *Proceedings of IEEE Infocom*, pages 104–113, 2001.

[2] A. R. Bharambe, M. Agrawal, and S. Seshan. Mercury: Supporting scalable multi-attribute range queries. In *Proceedings of ACM SIGCOMM*, August 2004.

[3] Census+ UI add-on for World of Warcraft. `http://www.warcraftrealms.com/censusplus.php`.

[4] C. Chambers, W. Feng, S. Sahu, and D. Saha. Measurement-based characterization of a collection of online games. In *Proceedings of the Internet Measurement Conference*, 2005.

[5] C. Diot and L. Gautier. A distributed architecture for multiplayer interactive applications on the internet. *IEEE Networks magazine*, 13(4), July/August 1999.

[6] EverQuest. `http://everquest.station.sony.com`.

[7] C. GauthierDickey, V. Lo, and D. Zappala. Using N-Trees for scalable event ordering in peer-to-peer games. In *Proceedings of NOSSDAV*, June 2005.

[8] C. GauthierDickey, D. Zappala, and V. Lo. A fully distributed architecture for massively multiplayer online games. In *Proceedings of ACM NetGames*, August 2004.

[9] J. Kim, J. Choi, D. Chang, T. Kwon, Y. Choi, and E. Yuk. Traffic Characteristics of a Massively Multiplayer Online Role Playing Game. In *Proceedings of ACM NetGames*, 2005.

[10] B. Knutsson, H. Lu, W. Xu, and B. Hopkins. Peer-to-peer support for massively multiplayer games. In *Proceedings of IEEE Infocom*, March 2004.

[11] D. Kushner. Engineering EverQuest. *IEEE Spectrum*, July 2005.

[12] The LUA Programming Language. `http://www.lua.org`.

[13] S. Rooney, D. Bauer, and R. Deydier. A federated peer-to-peer network game architecture. *IEEE Communications*, 42(5), May 2004.

[14] S. A. Tan, W. Lau, and A. Loh. Networked Game Mobility Model for First-Person Shooter Games. In *Proceedings of ACM NetGames*, 2005.

[15] World of Warcraft. `http://www.worldofwarcraft.com`.

[16] XPilot. `http://www.xpilot.org`.