Evaluating P2P Live Streaming Systems: the CNG Case*

Shakeel Ahmad[1] Christos Bouras[2] Eliya Buyukkaya[3] Raouf Hamzaoui[1]
Vaggelis Kapoulas[2] Andreas Papazois[2] Alex Shani[4] Gwendal Simon[3]

[1]Department of Engineering, De Montfort University, Leicester, UK
sahmad@dmu.ac.uk, rhamzaoui@dmu.ac.uk
[2]Computer Technology Institute & Press "Diophantus", Patras, Greece
bouras@cti.gr, kapoulas@cti.gr, papazois@ceid.upatras.gr
[3]Institut Telecom, Telecom Bretagne, France
eliya.buyukkaya@telecom-bretagne.eu, gwendal.simon@telecom-bretagne.eu
[4]Exent Technologies, Petach-Tikva, Israel
ashani@exent.com

Abstract

Many peer-to-peer (P2P) systems have been proposed for the provision of scalable live video streaming services over the Internet. While the literature contains surveys of the architectures of these systems, there is a lack of work on methodologies for their evaluation. We identify the main issues in the evaluation of P2P live streaming systems and use the Community Network Game (CNG) project as an example to illustrate them. The evaluation of the P2P system consists of two phases: a laboratory one using the ns-2 network simulator and an online field test with Massively Multiplayer Online Games (MMOG) players.

1. Introduction

Traditional client-server video streaming systems have critical issues of high cost and poor scalability. P2P networking exploits the upload bandwidth, computing power and storage space of the end users to reduce the burden on the servers and has been shown to be cost effective and easy to deploy. We identify the main issues in the evaluation of P2P live video streaming systems and use the CNG project as an example to illustrate them. The CNG project (`http://www.cng-project.eu/`) is an EU-funded research project that is focused on applying new network technologies to support community activities over highly interactive centrally managed MMOGs. CNG enhances collaborative activities between online gamers and develops new tools for the generation, distribution and insertion of User Generated Content (UGC) into existing MMOGs. It allows the addition of new engaging community services without changing the game code and without adding new processing or network loads to the MMOG servers. In particular, CNG proposes a P2P live video system to stream screen-captured video of MMOGs.

This paper presents the procedure that will be followed to evaluate the P2P live streaming system. Two phases are planned: a laboratory "offline" one using simulation software and an online one based on a real deployment. The online evaluation will be done by real gamers who will provide feedback through questionnaires. The performance of the system will be monitored by software, which will collect and provide useful information for further analysis.

The remainder of the paper is as follows. Section 2 provides an overview of the CNG P2P live streaming system. The plans for CNG laboratory experiments and online evaluation are presented in Section 3 and 4 respectively.

2. CNG P2P Live System

Allowing MMOG players to share their live game play with other players can have many useful applications. For example, skilled players can showcase their game to a large audience. Currently, the only platform that offers this service is Xfire (`http://www.xfire.com/`). Xfire captures the video of the game from the screen and sends it to a central server which broadcasts it live. However, this solution, which relies on central servers is expensive due to bandwidth and maintenance costs. To address these limitations, the CNG project proposes to use a P2P system. While many P2P live video systems have been developed, none of

*The research leading to these results has received funding from the European Commission's Seventh Framework Programme (FP7, 2007-2013) under the grant agreement no. ICT-248175 (CNG project).

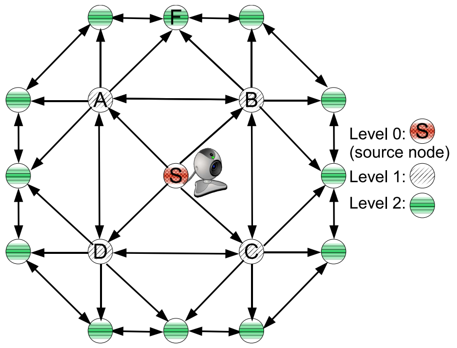

Figure 1. P2P topology.

them is suitable for the specific requirements of MMOGs:

- MMOG friendliness. The game experience should not be harmed by the P2P streaming. Thus, data communication with the MMOG game server must be given higher priority.

- Efficient management of multiple P2P overlays. Many MMOG players may simultaneously emit live streams, so the P2P overlay consists of many concurrent P2P overlays. A peer cannot participate in all P2P overlays because some of its resources will be used in every overlay it belongs to. The challenge for a user is to adequately allocate its physical resources, including upload and download bandwidth. These resources are limited, so they have to be shared carefully.

- Live video should be delivered at about the same time for all peers at the same "level". Moreover, peers in a higher level should in general be able to watch the video before those in a lower level. A level can be a priority class in a multi-tiered premium service. Alternatively, a level can be defined as the set of MMOG players that are in the same region of the virtual world.

In the following, we describe the CNG P2P system. The video of the game is captured in real time from the computer screen of the source and compressed. The resulting bitstream is partitioned into a sequence of source blocks, each of which corresponds to one Group of Pictures (GOP).

A mesh topology is used for the P2P network. This mesh is a directed graph $G = (V, E)$ where V is the set of peers, and (x, y) is in E if x may directly send packets to y. Peers are organized in levels. Level 0 contains the source. Level 1 consists of all peers that are direct successors of the source. In general, level k consists of all peers that are direct successors of level $k-1$ peers but are not in level $k-1$.

The UDP protocol is used as the transport protocol. The source applies rateless coding on each source block and sends the resulting encoded symbols in successive packets to level-1 peers until it receives an acknowledgment or a timeout occurs.

Packets are sent according to a scheduling strategy. The strategy specifies the maximum number of encoded packets n that can be sent by the source, the time at which a packet is sent, and a hierarchical forwarding scheme.

An example of a scheduling strategy for the P2P network of Fig. 1 is as follows. Packet 1 should be sent at time t_1 to A, which should forward it to B and D. Packet 2 should be sent at time t_2 to B, which should forward it to C. Packet 3 should be sent at time t_3 to C, which should forward it to B and D. Peer D should forward it further to A. Packet 4 should be sent at time t_4 to D, which should forward it to A and C. Peer C should forward it further to B.

Level-1 peers forward packets that are directly received from the source to their adjacent level-2 peers.

When a level-1 peer completes the decoding of a source block, it sends an acknowledgment to the source. Then it applies rateless coding on the decoded source block and starts acting as a source for level-2 peers that are its direct successors. The same procedure applies to peers at the next levels.

3. Laboratory Experiments

In the laboratory experiments, the CNG network solution with be evaluated with simulation software.

3.1. Metrics

To evaluate the performance of P2P video streaming systems, the following metrics were used in the literature: (1) Start-up delay: delay between the time a user joins a P2P system and the time it starts playing back the video [1] (2) Playback lag: time difference between the playback position of the source and that of the receiving peer [2] (3) Failure rate: probability that a user is rejected when it tries to join the system [1] (4) Continuity index: ratio of the number of video blocks that are available at their due playback time to the number of blocks that should have been played back by that time [1] (5) Peak Signal to Noise Ratio (PSNR) [3]. The PSNR is a standard video quality metric computed as: $PSNR(dB) = 10 \log_{10} \frac{255^2}{MSE}$ where MSE is the mean squared error between the original frame and the reconstructed frame (6) Percentage Degraded Video Duration (PDVD) [4]. The PDVD is the percentage of received frames whose PSNR is more than 2 dB worse than the PSNR of the corresponding encoded frames.

Both the PSNR and PDVD can easily be computed in the lab experiments as all required videos (original, encoded, received) are available.

In addition to measuring mean values of these metrics, we plan to divide CNG users into classes according to bandwidth, and measure minimum, maximum and variance values of the metrics for each class.

3.2. Examined Aspects

Scalability: Scalable systems are characterized by the property that the usage of resources is independent of the size of the system. Simulations for P2P live video streaming systems rarely consider more than several thousand simultaneous peers. Indeed, it is assumed that if the system scales at this stage, it is highly likely that it will scale to more peers. For the simulation of the CNG P2P video system, we plan to consider at least 2,000 peers.

Heterogeneity: Measurements show that P2P systems are characterized by a very high diversity of participating peers [5]. The variability of the *upload capacity* is considered as the major challenge in terms of heterogeneity, because it requires specific strategies in order to leverage the high capacity on some peers, and serve the peers with low capacity [1]. In [6], the average upload capacity of peers using BitTorrent has been found to be 180 kbps, while [7] shows an average upload capacity of 150 kbps. Two empirical distributions of peer upload capacities, one of broadband hosts (`http://www.dslreports.com/archive`) and one of BitTorrent hosts [8] are similar and are well modelled by a log-normal distribution [9]. Taking into account the above studies, the upload capacity in our experiments will follow a log-normal distribution with parameters μ in $\{150, 500, 1000\}$ and σ in $\{0.2, 0.9, 1.4\}$.

Churn: Based on the results from [10] and [7], we propose to use an exponential distribution of Time to Live (TTL) to model the churn rate. The parameter of this exponential distribution should result in approximately one tenth of the peer population refreshed at every simulation unit (which can be fixed at 1 hour). This setting, which is higher than what has been measured in [10], represents a system with a high churn. We will also evaluate our system under lower churn rates.

3.3. Simulation Tools

We considered several simulators including: ns-2, ns-3, P2PSim, Overlay Weaver, PeerSim, PlanetSim, Neurogrid, Query-Cycle, and Narses. We have assessed the available simulators based on several criteria, like simulator architecture, usability, scalability, statistics, underlying network simulation, and system limitations. We also considered the use of a testbed, like Planetlab. However, testbeds do not permit complete control of the system (e.g., for churn purposes) and, additionally, the scale of the experiments is severely limited. Taking into account the above, we decided to use the ns-2 simulator for CNG laboratory experiments.

3.4. Simulation Environment & Settings

A network game model for Counter-Strike is proposed in [11]. It is reported that 3–4% of all packets in a backbone could be associated with only 6 popular games [12]. One major concern is the upstream bandwidth. The peak percentage of traffic contributed by clients playing Counter-Strike to the total UDP traffic in the upstream direction can go up to 12% [13]. The work in [11] provides a simple traffic model for fast action multiplayer games. The game traffic model consists of only two independent modules, the client traffic model and the server traffic model with a burst size equal to the number of clients participating. In [14], the Extreme Value distribution has been identified to fit best for Quake traffic and other measurements have shown that newer MMOGs have bandwidth requirements that surpass those of older games [15].

It is important to identify user behavior with respect to network gaming. The study in [13] shows the percentages of subscribers playing Counter-Strike in various markets and identifies the trends in user behavior versus the day of the week. The common trend for all the markets is that this proportion increases as the weekend approaches, peaking on Friday, Saturday or Sunday. Also, there is a period of time in a day when there are very few to no users playing games. The exact hours vary depending on the market, indicating varying user behavior. The authors of [16] have filtered and analyzed MMOG traffic of the three-day long passive measurement, which contained about 200 World of Warcraft (WoW) flows and 100 other MMOG flows. Finally, [17] analyses a 1,356-million-packet trace from a sizable MMOG called ShenZhou Online.

3.5. Comparison to Similar Systems

Comparing the CNG system to state of the art commercial systems such as PPlive (`http://www.pptv.com/`) would present a major challenge as an implementation of these systems in ns-2 is not available. Moreover, the existing systems have not been designed for the CNG envisioned application. As an alternative, we propose to compare the CNG system to two related systems [18, 19] which although not designed for an MMOG environment can be implemented in reasonable time by making appropriate changes in the CNG system implementation.

4. Online Evaluation

In the online evaluation, MMOG players will be asked to answer an online questionnaire to describe their experience. In addition, the performance of the P2P network will be monitored using online tools.

Evaluation tools will be used for monitoring the activities and the traffic generated by the CNG system. Packet analyzers like Wireshark (`http://www.wireshark.org/`) are commonly used for traffic monitoring. In [20], Wireshark is used for analyzing and modeling the traffic generated by WoW. In [15], Wireshark was running on a computer along with Second Life, capturing game packets while filtering out irrelevant packets. On the other hand, built-in packet sniffers are useful for capturing the traffic and investigating protocols or applications in a non-intrusive way. When the source code of the application is available, traffic analysis can be done by adding proper logging and profiling functions to the source code [21].

In the online evaluation of the CNG network solution, Wireshark will be used to monitor the MMOG and other background traffic. On the other hand, the P2P network will be monitored through modules that will be developed in order to track user activities as well as the generated traffic.

For the online evaluation, a sample of 100-200 players will be recruited through forum posts and direct invitations to MMOG communities. An online user guide will be made available to help the players install and use the CNG tool. A number of evaluation measures will be used before, during and after the evaluation trial. The evaluations will be conducted online, allowing the project to obtain fast feedback. Participants will then be requested to continue their MMOG play and to augment it where they wish to with the CNG tools available to them. The usage period will be in the region of 4 weeks. Online players will be able to provide their feedback through a questionnaire where they will describe their experience with the CNG system, indicate any technical problem they encountered and suggest any further enhancement for the system.

5. Conclusion & Future Work

We have presented the plannning for the evaluation of the CNG P2P live streaming system. The process consists of two phases: laboratory experiments and online evaluation of the integrated system. The paper also includes a survey on the state of art for the evaluation of similar systems.

The next step of this work is the execution of the above plans. Our goal is that the execution of CNG evaluation will not only lead to significant results on the CNG system performance, but it can also introduce innovative ways of working for the evaluation of similar systems.

References

[1] B. Li, S. Xie, Y. Qu, G. Keung, C. Lin, J. Liu, and X. Zhang. Inside the new coolstreaming: Principles, measurements and performance implications. In *Proc. INFOCOM 2008*, pages 1031 –1039, Apr. 2008.

[2] D. Purandare and R. Guha. BEAM: An efficient peer to peer media streaming framework. In *Proc. Local Computer Networks '06*, pages 513–514, 2006.

[3] V. Padmanabhan, H. Wang, and P. Chou. Resilient peer-to-peer streaming. In *Proc. 11th IEEE International Conference on Network Protocols*, pages 16 – 27, Nov. 2003.

[4] 3GPP. Technical specification group services and system aspects; video codec performance (rel. 9), TR 26.902 v9.0.0. Technical report, 2009.

[5] J. Sacha. *Exploiting heterogeneity in peer-to-peer systems using gradient topologies*. PhD thesis, University of Dublin, Trinity College, 2009.

[6] O. Luzzatti. P2P live streaming for the masses - deployment and results. In *Proc. P2P-TV '07*, 2009.

[7] S. Khirman and V. K. Gurbani. Swarm analysis: 7 days of two popular swarms. In *Proc. P2PRG*, pages 26–31, Jul. 2009.

[8] M. Piatek, T. Isdal, T. Anderson, A. Krishnamurthy, and A. Venkataramani. Do Incentives Build Robustness in BitTorrent. In *Proc. NSDI'07*, 2007.

[9] A. Bharambe, J. R. Douceur, J. R. Lorch, T. Moscibroda, J. Pang, S. Seshan, and X. Zhuang. Donnybrook: enabling large-scale, high-speed, peer-to-peer games. In *Proc. SIGCOMM '08*, pages 389–400, 2008.

[10] H. Liang, R. N. Silva, W. T. Ooi, and M. Motani. Avatar mobility in user-created networked virtual worlds: measurements, analysis, and implications. *Multimedia Tools Appl.*, 45:163–190, Oct. 2009.

[11] J. Färber. Network game traffic modelling. In *Proc. NetGames '02*, pages 53–57, 2002.

[12] S. McCreary and K. Claffy. Trends in wide area ip traffic patterns - a view from ames internet exchange. In *ITC Specialist Seminar on IP Traffic Modeling, Measurement and Management*, pages 1–11, 2000.

[13] A. Sinha, K. Mitchell, and D. Medhi. Network game traffic: A broadband access perspective. *Comput. Netw.*, 49:71–83, Sep. 2005.

[14] M. S. Borella. Source models of network game traffic. *Computer Communications*, 23(4):403 – 410, 2000.

[15] J. Kinicki and M. Claypool. Traffic analysis of avatars in second life. In *Proc. NOSSDAV '08*, pages 69–74, 2008.

[16] G. Szabó, A. Veres, and S. Molnár. Effects of user behavior on mmorpg traffic. In *Proc. ICC'09*, pages 1531–1536, 2009.

[17] K.-T. Chen, P. Huang, C.-Y. Huang, and C.-L. Lei. Game traffic analysis: an MMORPG perspective. In *Proc. NOSSDAV '05*, pages 19–24, 2005.

[18] M. Grangetto, R. Gaeta, and M. Sereno. Reducing content distribution time in p2p based multicast using rateless codes. In *Proc. Italian Networking Workshop*, pages 1–12, Jan. 2009.

[19] C. Wu and B. Li. rStream: Resilient and optimal peer-to-peer streaming with rateless codes. *IEEE Trans. Parallel Distrib. Syst.*, 19:77–92, Jan. 2008.

[20] M. Suznjevic, M. Matijasevic, and O. Dobrijevic. Action specific massive multiplayer online role playing games traffic analysis: case study of world of warcraft. In *Proc. NetGames '08*, pages 106–107, 2008.

[21] W. Acosta and S. Chandra. Trace driven analysis of the long term evolution of gnutella peer-to-peer traffic. In *Proc. PAM'07*, pages 42–51, 2007.