Inside the Bird's Nest: Measurements of Large-Scale Live VoD from the 2008 Olympics*

Hao Yin[1], Xuening Liu[1], Feng Qiu[2], Ning Xia[3], Chuang Lin[1],
Hui Zhang[4], Vyas Sekar[4], Geyong Min[5]

[1]Tsinghua University, China [2]ChinaCache Co., Ltd, China
[3]Huazhong University of Science and Technology, China
[4]Carnegie Mellon University [5]University of Bradford, UK

ABSTRACT

The 2008 Beijing Olympics was an interesting event from a VoD perspective because it involved near real-time video delivery at massive scales over multiple days of a high-profile event. We present some measurement-driven insights into this event through a unique dataset obtained from ChinaCache, the largest CDN in China. The dataset is unique in three respects. First, it gives a "white-box" view into user access patterns which would otherwise be impossible. Second, since the CDN serves different content providers, it allows to compare and contrast the effects of different presentation models on end users. Third, the nature of the content itself is vastly different from traditional VoD systems in terms of the real-time and event-driven nature, which gives rise to unique effects. The dataset allows us to investigate a wide range of interesting issues: (1) how the live nature of the events causes differences in access patterns compared to traditional VoD and User-Generated Content (UGC) systems, (2) how the presentation models affect user behavior, and (3) flash-crowd phenomena. Based on these observations, we discuss implications for future live VoD systems.

Categories and Subject Descriptors

C.2.4 [**Distributed Systems**]: Distributed applications

General Terms

Measurement, Human Factors

Keywords

Video-on-Demand, Flash Crowds, User Behavior

*The Bird's Nest was the popular name given to the Olympics stadium in Beijing.

IMC'09, November 4–6, 2009, Chicago, Illinois, USA.

1. INTRODUCTION

The 2008 Beijing Olympics was an unprecedented event in many ways. We are specifically interested in this event from the perspective of Internet Video-on-Demand (VoD) systems. The Olympics saw a staggering scale of on-demand streaming never before seen on the Internet. In addition to the 3800 hours of live coverage, an additional 1200 hours of commentaries, highlights, and interviews were added, generating a total of 5000 hours [2] of online video content. In China alone, more than 80 million Internet users are believed to have accessed these videos [5].

The coverage spanned a three week long period and consisted of more than thirty thousand video clips triggered by thousands of real-time events. It contains a mix of both live and time-shifted videos and diverse presentation formats: full events, short highlights, interviews, on-line commentaries, social interactions, and advertisements models. In these respects, the video coverage of the Olympics differs significantly from VoD sites that offer TV episodes (e.g., hulu.com, abc.com, msn.com) or User-Generated Content (UGC) (e.g., YouTube). Given an event of this magnitude and importance, it is but natural to understand its implications on Internet VoD systems.

ChinaCache, China's largest Content Delivery Network (CDN) delivered most of the videos within China. (Due to copyright issues, ChinaCache only delivered video content for users in mainland China and Macau.) We were able to obtain traces capturing user access patterns during the entire duration of Olympics from ChinaCache. All the video sites used streaming delivery to deliver the videos and used Adobe Flash streaming technology to provide on-demand streaming. The dataset contains complete information of user actions: e.g., which videos they watched, when did they access these videos, their stop/pause/seek actions etc.

ChinaCache provided the back-end video delivery for three different content providers:

1. The official Olympic video web-site which was organized as a video portal (we label this *Off*).

2. The largest social networking site in China serving more than 530 million registered users (labeled *Soc*). *Soc* also used mechanisms to actively engage online users (e.g., using instant messaging).

3. The official syndicated video source site used by other web sites and web portals which provided five-minute highlights of the events (labeled *Synd*).

The "white-box" and unique nature of the measurements enables us to address questions that could not be answered otherwise. Since the measurements were obtained from the CDN, we can understand characteristics such as the time a video was first accessed and whether there are any hidden correlations between how different videos are accessed. Further, since our dataset is restricted to China mainland and Macau, which have a single time-zone, it allows us to study temporal effects and analyze events at multiple time granularities without worrying about timezone effects across the user population. These aspects become especially crucial for our in-depth understanding of flash-crowd phenomena (Section 5). Similarly, since the CDN uses streaming delivery, it provides us with detailed state information such as when the user hit play, pause, unpause, seek, and stop; this cannot be obtained with traditional HTTP delivery. Therefore, we can get the actual session times, after accounting for pause times, and thus reason about continuity in user viewing patterns. Finally, since we have three different views into how users obtained the videos, we can understand if the presentation models affected user behavior.

We answer several interesting questions using this dataset. We broadly classify these questions into one of three high-level categories:

- How does the live nature of the event impact the VoD system in terms of file sizes, file access popularity and end-user behavior? How are these different from traditional VoD and UGC websites?
- How do flash crowd like phenomena manifest in the context of such an event?
- How does the presentation model impact user behavior characteristics?

We present an in-depth analysis of flash-crowd phenomena [22, 24, 25] observed in our dataset. First, as a consequence of the way a live event was presented in near real-time as multiple smaller videos and as a result of hidden semantic relationships across videos, we observe correlated flash-crowds involving multiple related videos. Second, the social networking site contributes to flash crowds by actively prompting users to view certain videos. In fact, the flash crowds in the *Soc* dataset are noticeably larger than the other two datasets. Third, the flash crowds extend beyond just recently popular videos of a live event. Since multiple videos are related by content, similar or related videos generated in the past that are related to the live event also become popular.

Based on these observations, we derive some key implications for the design of future live VoD systems:

- **Caching:** Our analysis of user viewing time indicates that irrespective of the video duration, 80% of the users only view the first 10 minutes of a video. This means that caching mechanisms can benefit significantly by just caching initial segments instead of entire videos.
- **Streaming functionality:** Even though our infrastructure provided full video streaming functionality (e.g., allowing users to pause, seek, or rewind), we observe that majority of the users do not actually use such functionality, which is similar to the results in [20]. This suggests that VoD systems can implement simpler delivery modes (e.g., HTTP delivery) and still provide high user satisfaction.
- **Handling flash crowds:** The live VoD flash crowds we observe present a unique challenge as they can be unanticipated and often involve related videos including some released much earlier than the actual event. This suggests the need to design caching mechanisms that have a greater awareness of such relationships.
- **Presentation models impact user behavior:**
  1. The convergence of instance messaging and online video created unique effects in concentrating user accesses to fewer videos and also creating larger and faster flash crowds. As we see an increasing confluence of social networks, instant messaging, and online media (e.g., [6]), it becomes increasingly necessary to understand these effects better, not only from a commercial perspective but also for better system design.
  2. Users appeared to be more tolerant of advertisements embedded at the beginning of videos, when the videos had a real-time aspect to them, but non real-time videos with similar advertising had high attrition rates.

We present these implications with one caveat, however. The singular nature of the Olympics — high-profile, high-impact, global audience, live, multi-day etc. — means that some of these implications may be less applicable to commodity Internet VoD systems. By the same token, the singular nature magnifies the relevance of our measurement study and analysis to other high-impact events of a similar kind (e.g., World Cup Soccer, Euro 2012, Cricket World Cup) that are likely to exhibit effects that will not be captured by understanding commodity VoD systems.

The rest of the paper is organized as follows. Section 2 describes our measurement infrastructure and the access logs we collect. Section 3 gives a broad summary of the user access behavior, video file size distributions, and video access popularity and compares these with traditional VoD/UGC sites. Section 4 focuses on user behavior to understand how it was affected by video durations and streaming features. Section 5 provides an in-depth understanding of the flash-crowd behaviors we observed. In Section 6, we study how the presentation model affects user behavior. We discuss related work in Section 7 before concluding in Section 8.

2. MEASUREMENT OVERVIEW

2.1 Content Delivery Architecture

Figure 1 shows the VoD system deployed by ChinaCache. It consists of the data source, CDN nodes equipped with Adobe Flash Media Server (FMS), and the Operation Support System (OSS) responsible for load balancing, billing etc. The FMSes provide streaming delivery instead of traditional HTTP delivery (i.e., a simple download). This gives users functions to seek or pause videos.

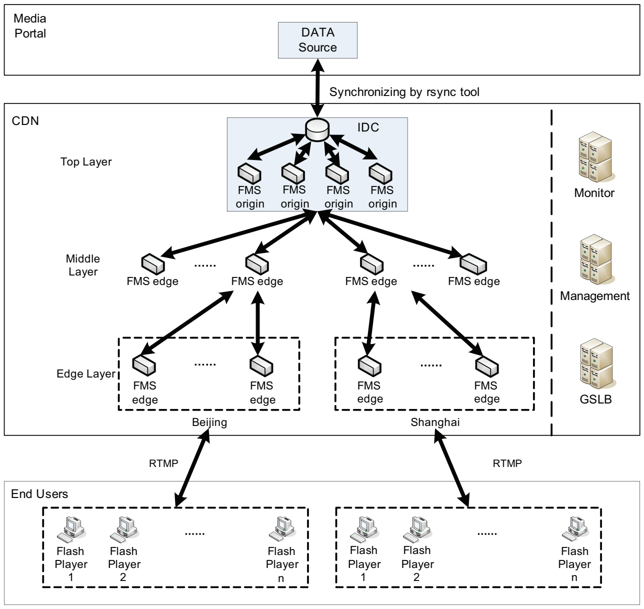


Figure 1: Architecture of the VoD system

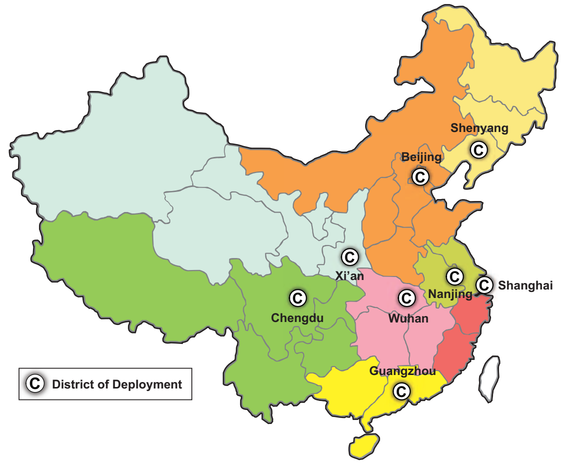


Figure 2: Deployment coverage of the ChinaCache VoD system

The FMS can be in either origin mode or edge mode. In origin mode, the FMS can obtain contents from local memory and storage, while in edge mode the FMS can request content from other FMSes. FMSes are organized into three layers. The top layer is composed of origin mode FMS servers and storage devices, on which video contents are periodically synchronized with the data source. For scalability and server capacity considerations, there is a middle layer FMS, which runs in edge mode. The edge layer FMSes directly serve end-users. Each FMS server is a Intel Dual Core Xeon 1.6GHz CPU with 4GB RAM with a network-attached storage system for disk I/O. The system uses a hierarchical caching scheme [4] to retain recently used files in the lower layers and less accessed files in the higher layers.

User requests are dispatched to closest edge layer FMS using traditional CDN techniques [3].[1] The end user sends a Real Time Messaging Protocol (RTMP) [1] request to the given edge layer FMS. This edge FMS serves the request if it has the content locally. Otherwise, the request is forwarded up the hierarchy of FMSes till it reaches some FMS that can serve the request.

2.2 Deployment

During the Olympics, the VoD system deployed by ChinaCache covered end users all over mainland China and Macau, as shown in Figure 2. More specifically, it included the users belonging to China Telecom and China Netcom, the two largest commercial ISPs in China, and the Chinese Education and Research Network (CERNET). For *Soc*, the system primarily serves users from Beijing and Shanghai. To cope with the large number of users, ChinaCache deployed 482 FMS servers in 8 districts (Figure 2). These districts are partitioned according to their geographic distribution and link quality. There are 27 top layer FMS servers. Around them, there are 95 FMS servers in the middle layer and 360 servers in the edge layer, which are scattered in each district. ChinaCache provisioned enough servers and bandwidth for the Olympics to face large access volumes (and flash crowds) and our access logs also confirm that the servers were not overloaded. Thus, our measurements are representative and not affected by server load.

[1]Due to copyright issues, ChinaCache only handled users from China; international users were redirected elsewhere due to the access restrictions.

2.3 Content Providers

While ChinaCache served as the backend content delivery platform for multiple websites during this period, we are specifically interested in the video requests originating from the three largest sources. These are the official Olympics video web site (*Off*), the largest social networking site in China (*Soc*), and the official Olympics video syndication site (*Synd*). Table 1 summarizes the relevant aspects of these three providers. These providers differed in two main aspects that we discuss below.

Business models: *Off* provided traditional video services for users from all over China. *Soc* not only provides portal services (e.g., news, entertainment reports, blogs) but also had special services for its own user group (e.g., instant messaging, online games). *Synd* was not directly released to the public. Rather, it was used by about 174 other websites around China on their front pages. Most of its users came from these websites. The three sources also differ in their advertising strategies too. Both *Synd* and *Off* embedded advertisements at the start of videos, while *Soc* did not.

Content presentation: *Off* provides a rich spectrum of videos with durations varying from a few seconds to a few hours. To provide the videos quickly, *Off* released the videos as soon as possible after the respective events, usually within 45 minutes. There are about 13700 videos including full-time games, highlights of events, interviews and on-line commentaries. *Soc* released around 20000 video contents similarly. *Synd* provided much less content compared to the other two. It served around 30 different highlights and fragments of the games played every day, each less than 5 minutes long. These videos were also replaced on a daily basis. As a portal,

	Off	*Soc*	*Synd*
Content Type	News, matches, highlights, interviews, commentaries etc.	News, matches, highlights, interviews, commentaries etc.	Mainly fragments of matches
Content Length	Varying from 10 seconds to 2.5 hours	Mostly in 5 minutes	Mostly 5 minutes
Content Resolution	480×360	400×300	320×240
Codes	On2 VP6 (Video) / MP3 (Audio)	Sorenson H.263 (Video) / MP3 (Audio)	On2 VP6 (Video) / MP3 (Audio)
Bit-rate	432 kbps CBR (Constant Bit-rate)	400 kbps CBR	332 kbps CBR
Content Indexing	Indexed by nations, matches, highlights, etc.	Indexed by nations, matches, highlights, etc.	No indexing
Advertising	2 pieces of embedded video advertisements	No embedded advertisements	2 pieces of embedded video advertisements
Navigation	From its own portal, and linked by the syndication site	From its own portal	Linked by about 174 web sites

Table 1: Comparison of the three content providers. *Off* is the official Olympics video website. *Soc* is the largest social networking site in China. *Synd* is the official Olympics video syndication site.

Off provided well-organized navigation capabilities for users by providing sitemaps, event categories, search features etc. The video player also provided thumbnails of related videos to guide viewers. *Soc* provided similar navigation features. Additionally, it also used client software to prompt and engage users (e.g., using instant messages). *Soc*, however, did not provide links to related videos. *Synd* was limited by the capabilities provided by the 174 web sites that linked to it; we do not have an extensive understanding of what presentation features these sites employed.

2.4 Data Collection and Processing

Each FMS server generates an access log that tracks the behaviors of server programs and end users. We collect the log files from all edge FMSes between Aug. 6th to Aug. 31th, 2008, spanning the duration of the Olympic Games. Each FMS server periodically reports a log every 15 minutes to the OSS and the uploaded logs are then merged on a daily basis.

Each entry in the access log represents different server and user actions. For example, these include server actions like server-start, server-stop and user actions (e.g., connect, disconnect, play, pause, unpause, seek, stop). Our focus in this paper is primarily on the user actions. We focus on six relevant types of fields from the access log described in Table 2. During this period, more than 3.5 million unique clients accessed over 34,000 video files in the system generating around 170 million lines of full log traces from the edge FMS servers.

We analyze the events in the access logs to capture the process of a user accessing a video. We refer to a set of events having the same client IP and client ID fields as a *connection*. The process of a user accessing a video might span multiple events in the log. To model user actions within a connection, we developed a simple state machine model to capture user actions (see Figure 3).[2] In the simplest case, it involves five actions in order: connect-pending, connect, play, stop, disconnect. These actions are associated by the common connection identifier of the user which is specified by client IP and ID fields in the log. We further split a connection into *sessions* (see Figure 4). Each session starts with the action "play" and ends with a "stop" action with a number of operations between these.

[2]Some connections have undefined transitions; this leads to an error state and we discard these entries. More than 97% connections were valid.

#	Category	Description
1	Action	Action name, such as connect, play, and status code, such as "200" which means success, "404" which means a file-not-found failure.
2	Time	Date and time when this action occurred, and also the time duration of this action.
3	Server	FMS server's IP address, load, and process information.
4	User	User IP address, client ID (a unique number generated on a FMS server for different client connections), user agent, referrer URL.
5	Content	Content's URL (include the domain name for publishing this content), protocol, file name, length, size in bytes, and accessed position in milliseconds.
6	Traffic	Network traffic from server to client, and also from client to server.

Table 2: Description of different types of data collected in the access logs

2.5 Summary Statistics

Table 3 summarizes our access logs. More than 3.5 million unique clients established around 18 million connections and accessed over 34,000 video files in 20.4 million sessions. The VoD system used dedicated FMS servers for each provider. Thus, we can classify the events according to the three different content providers by simply using the edge FMS servers used to serve the specific content. Note that since all the users and servers are located in the same time zone, we have a completely synchronized view of the events. This is especially useful as it allows to study different temporal patterns and analyze the data at different time granularities (e.g., per-minute, hourly, daily).

3. UNDERSTANDING LIVE VOD

In this section, we present the broad characteristics of the Live VoD measurements and qualitatively compare these with measurements reported from traditional VoD and UGC

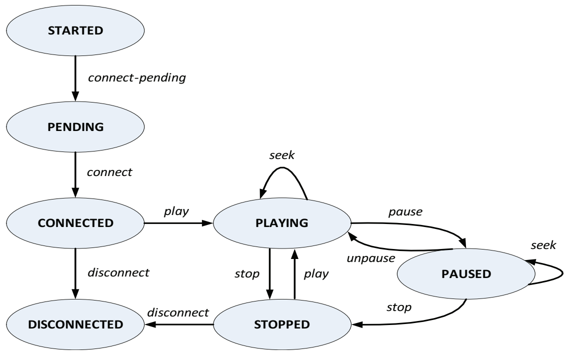


Figure 3: State machine for a connection process

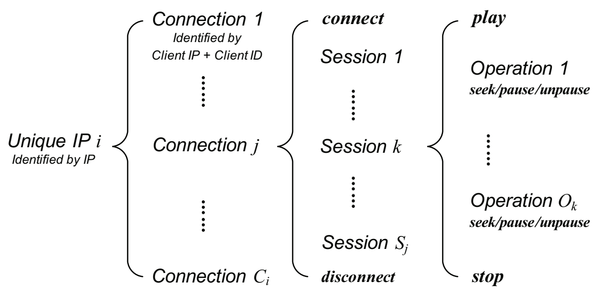


Figure 4: Illustrating the definition of a connection and session

deployments. We refer to our workload as Live VoD, in the sense that the content is changing in real-time similar to a live streaming workload and at the same time the videos are stored to be served later on-demand, this is similar to a VoD workload. In particular, the real-time, event-driven nature of the Olympics gave rise to unique patterns in the distribution of video lengths (Section 3.1), temporal access patterns (Section 3.2), and evolution in the popularity of content over time (Section 3.3).

3.1 Video Lengths

In Live VoD, the video length is very diverse, ranging from a few seconds to a few hours. This was an effect of (1) the nature of the actual events: different sports contests spanning different durations (e.g., swimming, volleyball, archery), press conferences, award ceremonies etc., and (2) the segmentation of a longer event into multiple smaller videos so that the content providers could give their users a *near real-time* feed. Traditional VoD systems see far less diversity in video durations. For example, in the case of VoD for distributing TV episodes or movies, the video lengths show a clear concentration in specific ranges: between 30–60 minutes for TV episodes, or 90–120 minutes for movies [28]. Similarly, in UGC sites (e.g., YouTube), 98% of the videos are less than 600s [14].

Figure 5 shows the CDF of the video lengths in the three content providers. For *Synd*, the CDF shows a sharp peak at 300 seconds. For *Off*, 70% of the videos are less than 1000 seconds. The linear pattern in *Off* between 70% to 100%

Content Provider	# Log entries (millions)	# Unique IPs (millions)	# Connxns / Sessions (millions)	# Videos accessed (thousands)
Soc	118	2.5	14/14.2	20
Off	47	1	4/5.8	13.7
Synd	5	0.12	0.15/0.4	0.05

Table 3: Summary statistics across the three different content providers

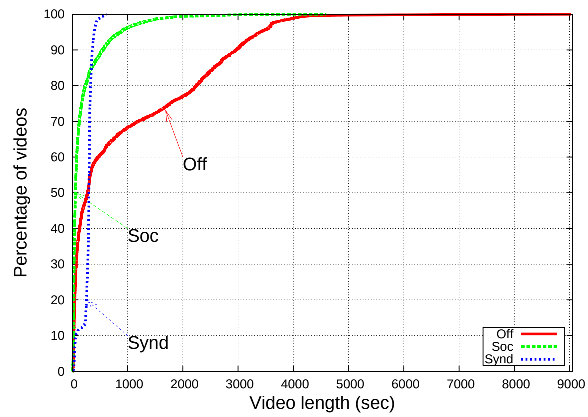


Figure 5: CDF of video length

suggests a uniform distribution of longer videos. *Soc* has much shorter videos; the 80th percentile of video duration is around 300 seconds.

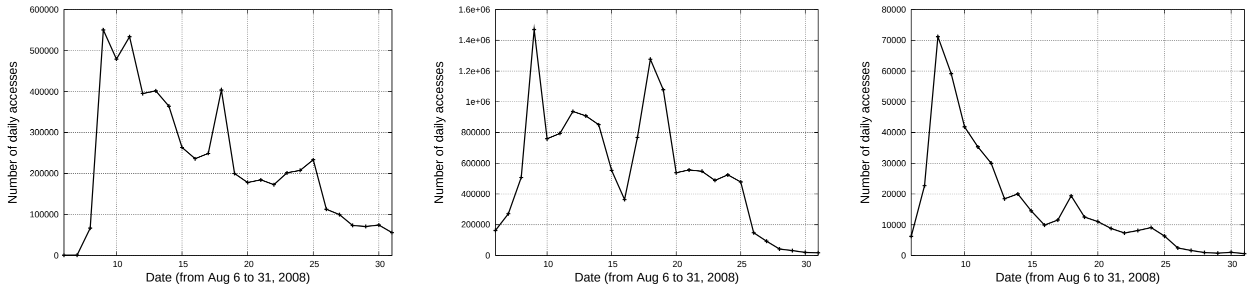
Most videos in *Synd* are edited to 300 seconds and published significantly after a live event occurred. Thus it does not capture the real-time, event-driven features of live VoD. As such, for the following results, we do not include results from *Synd*.

3.2 Temporal Access Patterns

Figure 6 shows the total number of videos accessed per day over the 24 day period from August 8 to August 31. While the number of accesses across the three providers is different, we see a similar pattern: a peak on Aug 8/9 with a gradual decrease subsequently. There are also two obvious peaks on Aug 8/9 and Aug 18. Aug 8 was the opening ceremony; as expected a large number of viewers tuned in to view it. Aug 18 is particularly interesting; an injury forced the popular Chinese athlete Liu Xiang to withdraw from the 110-meter hurdles in the first round [7]. While the Aug 8 peak was anticipated, the Aug 18 peak certainly was not.

Looking at the number of accesses per hour in Figures 7(a) and 7(b), we see some time-of-day effects, with the average traffic volume peaking around 10 pm every day (ignoring the large spikes in the graph). The time series plots clearly show sharp peaks in certain hours indicative of flash crowd like behavior. For example, the number of accesses spikes to 160,000 for *Soc*; this is five times the average number of accesses per hour.

For each day, we identify the *rush hour* – the hour of day that attracts the most accesses within the day – and plot the proportion of daily accesses contributed by the rush hour in Figure 8. We see that the rush hour attracts a significant



(a) *Off* (b) *Soc* (c) *Synd*

Figure 6: Daily accesses (covering the Olympics in Aug 2008)

portion of the total number of accesses with each day. For example, in *Soc*, this can be as high as 17.6%, on Aug 13 12:00 pm.

Measurements of traditional VoD systems have shown that the rush hour effect exhibits time-of-day effects [21, 23, 28]. Figure 9 shows for each day, the actual hour of day during which the rush hour occurred. However, as Figure 9 shows, there are no discernible time-of-day effects in the distribution of when rush hours occur. Further, we can visually confirm that the rush hours of *Soc* and *Off* are strongly correlated. For 8 out of 17 days, the rush hours of these came within one hour of each other. These results suggests that a flash-crowd effect might be occurring. We revisit these when we analyze flash crowds in-depth (Section 5).

3.3 Video Popularity in Live VoD

3.3.1 Static Popularity Patterns

Many real-world phenomena including traditional VoD and web popularity exhibit the Pareto principle (80-20 rule) [11, 28]. We study if similar behavior manifests in live VoD as well. To analyze this, we identify for each day, the top- k% (k=10,20) of the videos sorted by access counts. Then, we compute the fraction of total accesses for that day contributed by these top-k% of the videos.

Figures 10(a) and 10(b) show the percentage of accesses contributed by the top-10% and top-20% of the videos for *Off* and *Soc*. In the period spanning the actual games (Aug 8–24), we can see that the top 10% and top 20% of videos contribute 80% and 90% of the total accesses for *Off*. As Table 3 shows, *Soc* had far more viewers than *Off*. Interestingly, with more users the popularity of videos tends to become more skewed; the top 10% and top 20% videos attracted 90% and 95% of the total accesses. We posit that this is a consequence of the presentation model in *Soc* and explore this further in Section 6.

3.3.2 Popularity Dynamics

An interesting question is how does the popularity of content change across time in such a multi-day VoD event? Does a specific video remain "hot" over multiple days or will it be superseded by videos pertinent to more recent events?

To analyze changes in the pattern of popular videos, we compute the set of top-k (e.g., k=5, 10, 20, 100) videos and understand how this set changes daily. To quantify the churn (i.e., how dynamically does the popular content change), we define the change-per-day percentage (CDP) across the top-k videos (for k=5, 10, 20, 100). Let P_i^k denote the set of top-k popular videos on day i. Then the top-k change per day percentage for day i, $CDP_i(k)$, is defined as:

$$CDP_i(k) = \frac{|P_i^k - P_{i-1}^k|}{k} \times 100\% \qquad (1)$$

Figures 11(a) and 11(b) show the $CDP(5)$, $CDP(10)$, $CDP(20)$, and $CDP(100)$,from August 7 to August 31, for *Off* and *Soc*. From these results, we make three main observations:

1. The popular content changes frequently. Between Aug 8 and Aug 25, $CDP(5)$ is almost 100%, which means that the top-5 videos were completely new. These hold for $CDP(10)$ and $CDP(20)$ as well. Even $CDP(100)$ is 55% for *Off* and 70% for *Soc*. Measurements from traditional VoD systems have typically reported values in the range of 16% to 30% for $CDP(10)$ and 12% to 16% for $CDP(100)$ [28]. This suggests that the real-time, event-driven nature results in the popular content changing more dynamically.

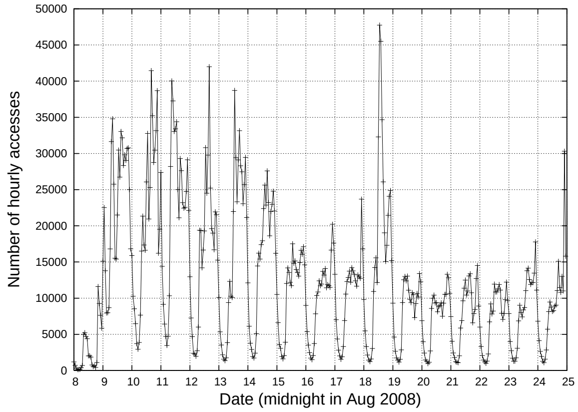
2. Certain events can keep users attention for a long time. For example, the churn in popular content is quite low on August 9 and August 19. Recall that these were immediately following two main events: the opening ceremony and Liu Xiang's withdrawal. These events, one anticipated and one unexpected, attracted the attention of a large number quantity of viewers.

3. We see an immediate and significant drop in the CDP values in the post-games period. The videos published during this time were mostly retrospective videos with little real-time impact and thus do not attract many users. In fact, in this period, the CDP values almost revert to those observed in traditional VoD systems.

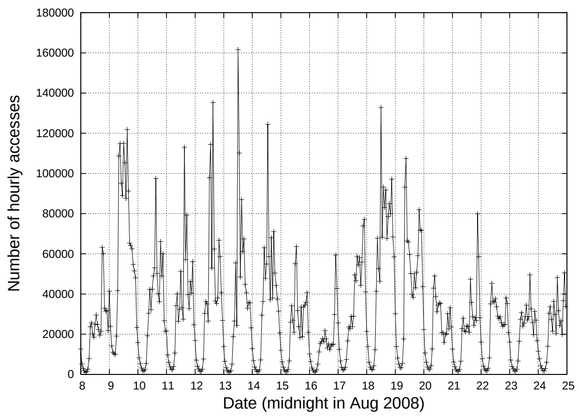
3.4 Observations

The nature of the Olympics workload – real-time, event-driven videos means that the durations, access patterns, and popularity dynamics are significantly different from traditional VoD and UGC systems. From the above results, we highlight three key observations:

1. Video lengths show significantly more diversity compared to traditional VoD systems. A natural question therefore is if user behavior was affected by the video lengths. We analyze this in the next section.



(a) *Off*



(b) *Soc*

Figure 7: Number of access per hour

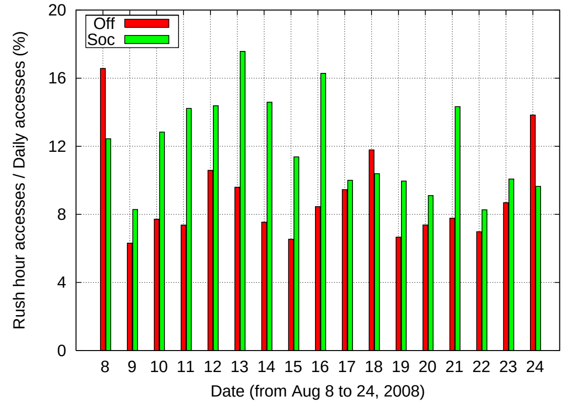


Figure 8: Fraction of total accesses contributed by the "rush hour" each day

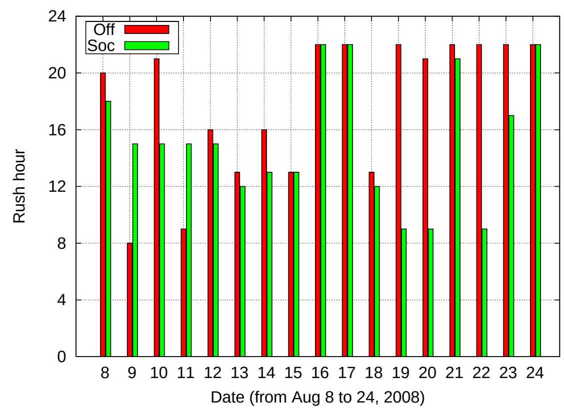


Figure 9: When did the rush hours occur?

2. There are noticeable flash crowd like effects in live VoD which are not typical of traditional VoD systems. We present case studies to better understand flash-crowd effects in Section 5.

3. Popular content changes significantly more rapidly than traditional VoD. This is a natural consequence of the event-driven nature of the content; recent events supersede past events in popularity.

4. Additionally, we observe that *Soc* shows more skew in content popularity. This suggest that the presentation model can affect access patterns; we explore this further in Section 6.

4. UNDERSTANDING USER BEHAVIOR

There are two features that distinguish the Olympics VoD system from traditional VoD and UGC systems. First, the system features a much wider distribution of video durations. Second, the system provides users with streaming capabilities (e.g., seek, pause) operations.[3] In this section, we understand how these impact user access patterns.

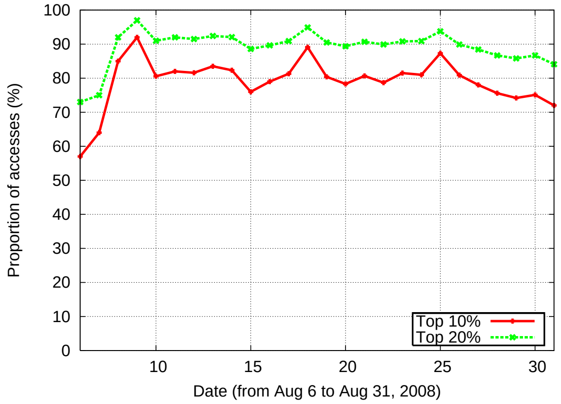
[3]Earlier versions of UGC sites did not provide these capabilities. Recently, they have retrofitted these features to traditional HTTP delivery.

4.1 Impact of Video Length on User Behavior

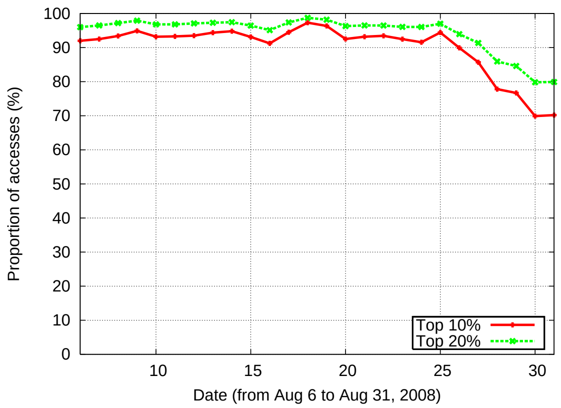
We define the viewing percentage as the fraction of a video duration taken up by each session (after accounting for pause intervals). For example, if the video was 100 seconds long and the session lasted 75 seconds, the viewing percentage would be 75%. (There is a possibility that the user replayed some earlier portion of the video. This might result in overestimating the viewing percentage. However, as we show in Section 4.3, most sessions had no seek actions. Thus, the viewing percentages are accurate.) We select five distinct video lengths: 60, 120, 300, 602 and 1812 seconds, and plot the CDF of the viewing percentage of these five classes in Figure 12. A general trend is that the viewing percentage is inversely proportional to the video duration. For example, for 120-second videos, more than 80% sessions have viewed at least 80% of the video. For the 1812-second videos, 70% of the sessions viewed less than 20% of the video.

4.2 Session Times

Figure 13 shows the distribution of session times for the three providers. We see that more than 80% of the sessions have a viewing time under 600 seconds across all three providers. This is surprising, considering that *Off* had significant diversity in video durations. To understand this further, we show in Figure 14 the viewing duration as a function

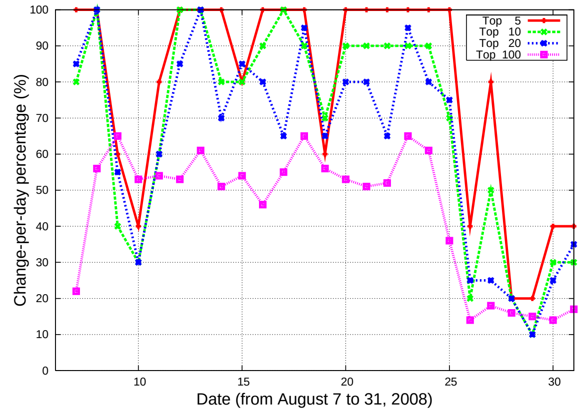


(a) *Off*

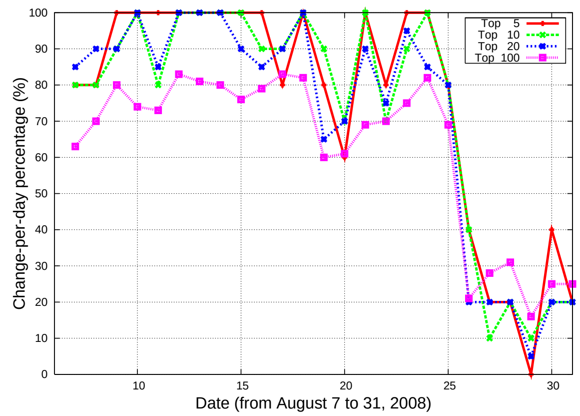


(b) *Soc*

Figure 10: Understanding the skew in number of accesses per video: contribution of the top 10% and top 20% videos to total accesses



(a) *Off*



(b) *Soc*

Figure 11: Change-per-day percentage of top 5, 10, 20, and 100 videos

of video length. For each video length, it shows the average, min, and max values across all sessions that accessed videos of that length. We see that for low video lengths, the viewing time is strongly correlated with the length. However, as the video length increases, this correlation weakens significantly. For example, 96% of the videos longer than 20 minutes have an average viewing duration less than under 400 seconds. 73% of such videos in fact only have a viewing duration between 200 and 400 seconds.

4.3 Do Viewers Use Streaming Capabilities?

For each session, we count the number of user actions between the "play" and "stop" events. Surprisingly, we find that around 80% of the sessions have *no* user operations (i.e., no pause or seek behavior). This effect has also been reported in previous measurement s studies [20], in which the percentage of sessions without interactivity were between 60% to 80%. This result is relevant for two reasons. First, even though users had these capabilities, they chose to not exercise these functions. Second, for the earlier results, we defined the viewing duration as the length of the session: the time from the user hit play to the user pressing stop and subtracting any intervals of pause action. However, it is possible that there were seek operations; this could mean that some parts of the video might have been replayed. Since we find that most videos had no seek operations, our definition of viewing duration is still valid.

4.4 Implications of User Behavior

There are two interesting observations from our study of user behavior: (1) most sessions last less than 600 seconds irrespective of the actual video duration and (2) users did not exercise the streaming functions to seek or pause videos. We derive two implications from these observations:

1. The vast majority of viewers only see the first 600 seconds of most videos. This can guide the design of more effective caching mechanisms – it is better to cache the initial segments of many videos instead of caching large videos in their entirety.

2. Since most users do not appear to use the streaming capabilities, simpler delivery modes (e.g., HTTP delivery as used by current VoD mechanisms) might suffice, without adversely affecting the users' viewing experience. Some UGC sites have modified their delivery techniques to retrofit streaming functionality; it would be interesting to see if and how users utilize these capabilities.

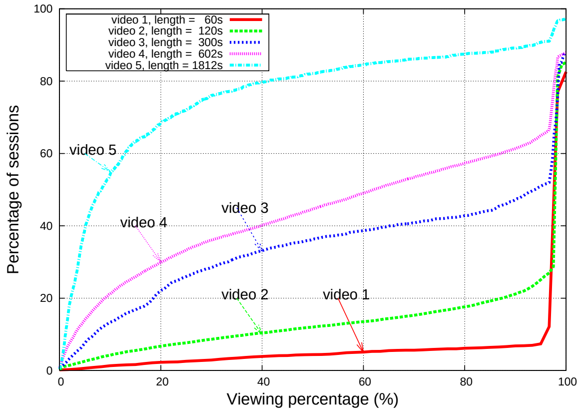


Figure 12: CDF of viewing percentage of different sessions

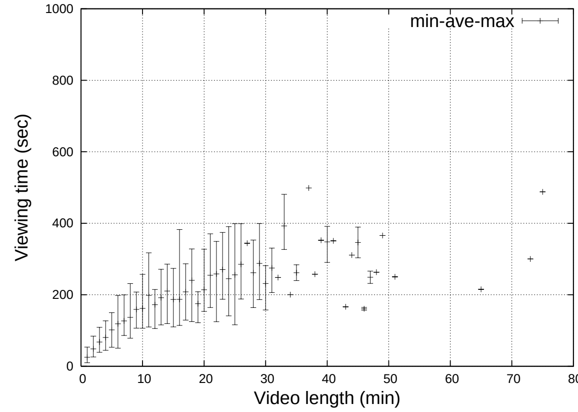


Figure 14: Correlating viewing duration with video lengths (grouped per minute)

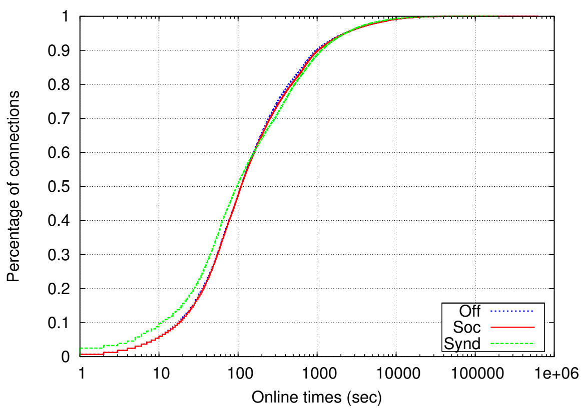


Figure 13: Distribution of session times across different providers

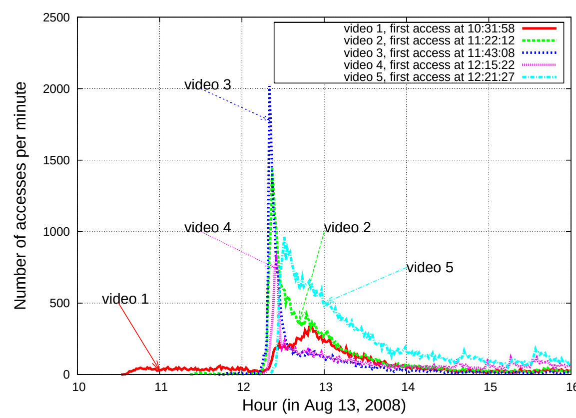


Figure 15: Number of accesses per minute for the top 5 videos in *Soc* on Aug 13 12:00 pm (on-going videos published during the game)

5. ANALYSIS OF FLASH CROWDS

Traditional VoD or UGC systems do not exhibit flash crowd like phenomena. However, in the earlier measurement results, the distinct spikes in the number of accesses per hour (Figures 7(a) and 7(b)), the large contribution of accesses by the rush hour (Figure 8), the rapidly changing profile of popular videos (Figure 11), and the strong correlation in when the rush hour occurred across *Soc* and *Off* (Figure 9), together strongly indicate flash-crowd effects.

In this section, we present a more in-depth understanding of flash-crowd phenomena in the live VoD context. First, we correlate the observed large flash crowds with actual events during the Olympics to confirm that there was indeed such flash-crowd effect. Next, we take three flash crowds as case studies to understand (a) if the flash crowd spanned multiple videos and (b) the effect of publish time on videos constituting a flash crowd.

5.1 What Triggered Flash Crowds?

To understand which specific events triggered flash crowds, we look at the rush hours with the largest number of accesses from Figures 8 and 9. For each rush hour, we identify the top-10 videos and manually find the most common real-world event that relates videos within this set. Using this approach, we annotate each flash crowd. Table 4 shows the events identified from *Soc* and *Off*. Unsurprisingly, given that the users in our dataset are based in China, we see a bias toward events involving popular Chinese athletes or team events.

5.2 Correlated Accesses in a Flash Crowd

Next, we analyze one specific flash crowd that occurred at 12:00 pm on Aug 13. We picked this specifically because this was the rush hour with the most accesses in *Soc* and the entire dataset (Figure 7). We selected the top 5 videos in terms of number of accesses and found that all videos were for the same event: Chinese women gymnasts won their first ever Olympic team gold medal.

Figure 15 shows the number of accesses per-minute for these 5 videos. The earliest video, video 1, was released at 10:31, and videos 2 to 5 were released at 11:22, 11:43, 12:15 and 12:21. Each video was a segment of the actual gymnastics event and video 3 was the segment in which the Chinese team actually won the gold medal. We see that the early videos continue to attract a large number of uses till 12:25. In this case, the correlations arise naturally since all the videos are segments of the same logical event.

Date	Events		Rush Hour		% Contribution to daily accesses	
	Off	*Soc*	*Off*	*Soc*	*Off*	*Soc*
Aug.8	Opening ceremony	Opening ceremony	20:00	18:00	16.6%	12.4%
Aug.12	Diving: Women's Synchronized Platform	Diving: Women's Synchronized Platform	16:00	15:00	10.6%	14.4%
Aug.13	Gymnastics: Women's Team Competition	Gymnastics: Women's Team Competition	13:00	12:00	9.6%	17.6%
Aug.16	Basketball: men's China vs Germany	Athletics: Men's 100m	22:00	22:00	8.5%	16.3%
Aug.17	Shooting: Men's 50m Rifle bronze	Diving: Women's Individual Springboard	22:00	22:00	9.5%	10.0%
Aug.18	Liu Xiang's withdrawal	Liu Xiang's withdrawal	13:00	12:00	11.8%	10.4%
Aug.24	Closing ceremony	Closing ceremony	22:00	22:00	13.8%	9.6%

Table 4: Specific events in the Olympics that triggered large flash crowds

Intrigued by the fact that the top-5 videos for this particular event were correlated, we proceed to see if similar correlations appeared in the remaining rush hour periods as well. We pick out the top-10 videos with the most accesses in each rush hour and check if these were related as well.

To define if two videos are "related" we need to understand hidden semantic relationships between videos. We make a simplifying assumption and define a narrower relationship. Videos that pertained to the same "sporting event" in some way: the actual competition event, prize ceremony, reports, and interviews are labeled as "related". (This could be underestimating the correlation since we do not look for latent semantic relationships. We present a case study on semantic relationships later.) We identify such correlations by manually inspecting the videos in these top-10 sets.

We take the top-10 videos for the rush hour. For each video, we find other videos in this set that are related (as defined above). Let $CS(v)$ denote the correlation set for video v in the top-10 set. Let $\hat{v}$ denote the video with the largest correlation set in this rush hour. To understand how correlated the top-10 videos are, we look at two measures: (1) The relative size of the largest correlation set, $RelCS_{max} = \frac{CS(\hat{v})}{10}$, and (2) The relative contribution of accesses from this largest correlation set $RelAcc_{max} = \frac{\sum_{v \in CS(\hat{v})} Acc(v)}{\sum_{v \in Top10} Acc(v)}$, where $Acc(v)$ is the number of accesses for video v during the rush hour.

Figure 16(a) and Figure 16(b) show the $RelCS_{max}$ and $RelAcc_{max}$ for *Soc* and *Off* respectively. We see that on Aug 13 (*Soc*) and Aug 8, 9, 13, 16 (*Off*) the top 10 videos of each rush hour are all correlated. This further corroborates that these events were flash crowds – a large number of accesses are triggered by the same event. For *Off*, 11 out of 17 days exhibits strong correlation at rush hour with $RelCS_{max}$ greater than 50%. Similarly, 8 out of 17 days for *Soc* have $RelCS_{max}$ greater than 50%. In other words, the set of top-10 videos were strongly related. Further, in both *Off* and *Soc*, 7 out of 17 days have $RelAcc_{max}$ greater than 70%; i.e., the accesses contributed by the correlated set dominate the total number of accesses.

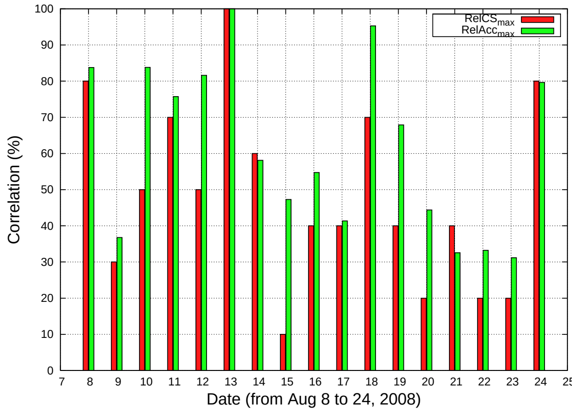
As mentioned earlier, our definition of "related" videos is quite narrow. However, we do observe some cases where the flash crowd had other semantically related videos. (Unfortunately, these are harder to quantify.) However, as an example, we consider the rush hour with second largest proportion of accesses in *Soc* (22:00 on August 16, in Table 4). This event was the final of the men's 100 meters (in which Usain Bolt set a new world record, and needless to say, won in style with a fair distance to spare!). This event triggered a large number of accesses for the preliminary rounds of the 100m event that occurred earlier. The access patterns for these videos is shown in Figure 17. Videos 1 and 2 are for the qualifying stage and had been published over 24 hours earlier. However these videos also attracted flash crowds along with the flash crowd for video 3 (the actual final race).

5.3 Does Publish Time Matter?

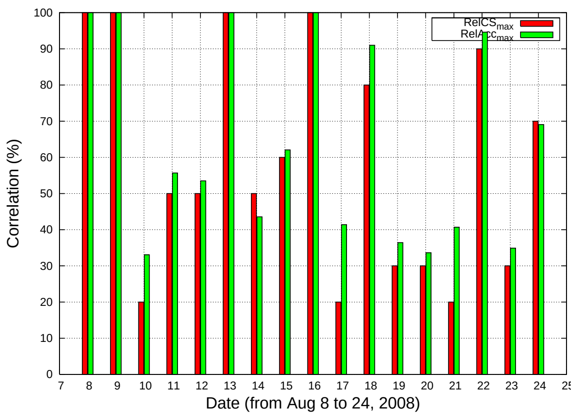
Given the real-time, event-driven nature of live VoD, we want to understand if the access pattern has any relationship with the publish time. To examine this relation, we analyze the rush hour with the most number of accesses in *Off*. This event was the press conference held after Liu Xiang withdrew from the games. Even as the press conference was going on, videos were being created and published. We select four videos from this rush hour which pertain to the same event.

Figure 18 shows the number of accesses per minute as a function of time (specified relative to the first access time). The earliest video was first accessed at 12:35:44 pm that day. At the time this video was released, the press conference had not ended. The number of accesses increases sharply as soon as the video was released. Among the four videos, this earliest video has the largest peak access rate (more than 200 accesses per minute). Compared to other videos, the first video also has the sharpest increase, i.e., for the access rate to reach the peak value. We refer the time duration between the first access to the time when the access rate reaches the maximum as the *time-to-peak*. The next two videos have their first access at 12:55:49 pm and 13:04:36. The peak access rates of these two videos are 50% lower than the earliest video. The last video is first accessed at 14:24:24. Although it is actually the most complete version compared to the others, its peak access rate is the lowest.

Note that this series of videos is slightly different from the last section. In the previous set, the key event was video 3, when the Chinese gymnasts won the gold medal – this was the video with the largest volume in the flash crowd. However, in the case of the press conference, the actual event (Liu Xiang's withdrawal) occurred much earlier, so the start of the press conference (video 1) was the key event in the series. In both cases, the appearance of a flash crowd is directly related to the release time of the most anticipated video. Further, in this case, the sooner a video was released, its chances of resulting in a flash crowd were greater.



(a) *Soc*



(b) *Off*

Figure 16: Correlation in rush hour videos

5.4 Implications of Flash Crowd Analysis

1. It is clear that flash crowds bring a huge volume of requests. On one hand, such events are hard to forecast, and it is difficult for service providers to anticipate them. On the other hand, since these flash crowds involve several related videos, users can be satisfied by providing related videos – either earlier time segments of the same event or other semantically linked videos – even if the newest video is not immediately accessible. Thus, a content provider can push such related content to different edge servers during such flash crowds.

2. The live, real-time nature of the event suggests that early release attracts more traffic. Thus, it might be prudent for the VoD system to *defer* releasing new content during overload situations.

6. IMPACT OF PRESENTATION MODELS

In this section, we analyze the impact of different presentation models. Specifically, we try to understand if there were significant differences between the three presentation models in terms of content popularity, flash-crowd behaviors, and access patterns. We also measure the impact of pre-video advertisements on user interest.

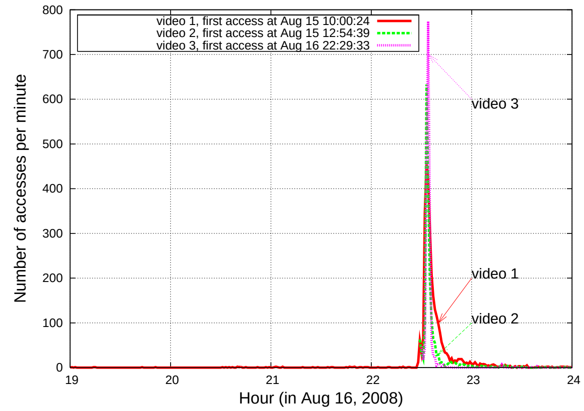


Figure 17: Flash crowd during the men's 100m final with related videos released much earlier

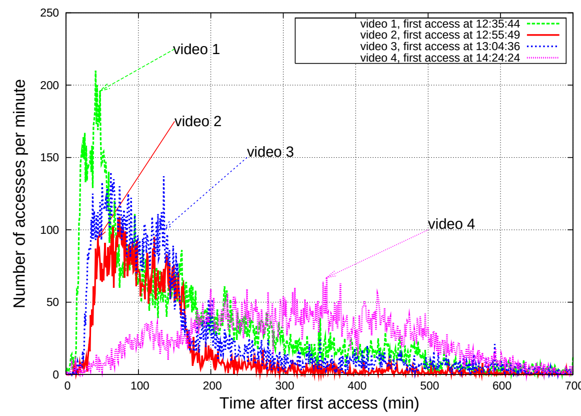


Figure 18: Number of accesses per minute in *Off* for the press conference videos following Liu Xiang's withdrawal on Aug 18

6.1 Access Concentration

We define the time-to-peak for a flash crowd as the time elapsed between the time a video was first accessed to the time the access rate (measured in number of accesses per minute) reaches its peak value. Table 5 compares the time-to-peak values for the different flash-crowd phenomena shown in Table 4. We see that *Soc* brings much faster access (i.e., shorter time-to-peak) value. Reflecting back on the rush hour graph (Figure 9), we see that the rush hour for *Off* typically lags *Soc* by one hour – this is a result of the faster accesses brought by *Soc*.

We also compare the 80% percentile period. This is defined as the length of the time taken for a video to reach 80% of its total accesses. Figure 19 shows the CDF of 80% percentile period for *Soc* and *Off* for the 200 most popular videos. It takes around 8,000 minutes for 80% of videos to reach their 80% percentile in *Off*, while it takes only 1,200 minutes in *Soc*. In other words, *Soc* brings more concentrated accesses to its videos.

Date	Events	Time-to-Peak (min)		80% percentile period (min)	
		Off	*Soc*	*Off*	*Soc*
Aug.10	Diving: Women's Synchronized Springboard	38	12	1140	1113
Aug.12	Diving: Women's Synchronized Platform	59	16	614	58
Aug.13	Gymnastics: Women's Team Competition	39	10	461	562
Aug.14	Gymnastics: Men's Individual Competition	40	9	2028	369
Aug.15	Athletics: Men's 100m	87	5	6384	12
Aug.18	Liu Xiang's withdrawal	41	14	3154	95

Table 5: Time-to-peak and 80% percentile period

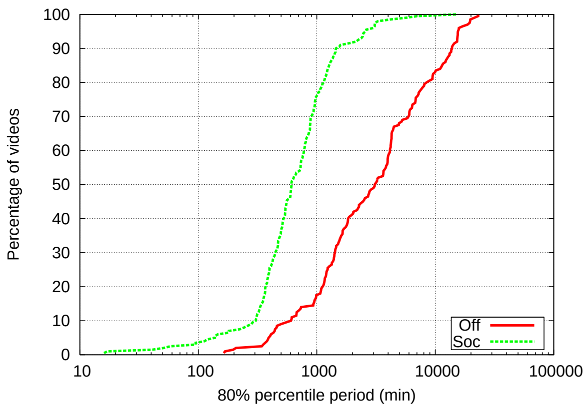


Figure 19: CDF of 80% percentile period across *Off* and *Soc* for their top 200 hot videos

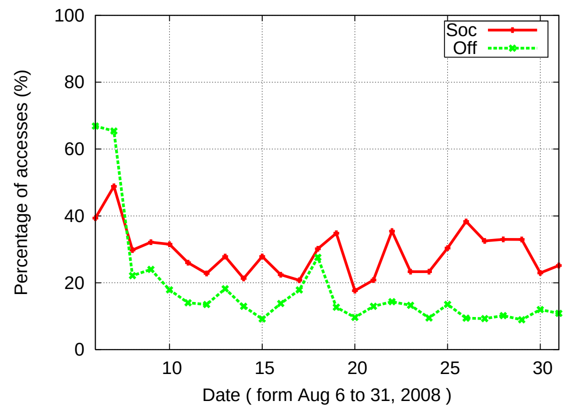


Figure 20: Fraction of accesses for top-5 videos

6.2 Video Popularity

We already saw that video popularity exhibits a Pareto-like principle. To understand if there is a significant difference between presentation models, we analyze the contribution of the top-5 videos per day to the total number of accesses. Figure 20 shows that overall *Soc* has a more skewed distribution – the top-5 videos contribute 29% of accesses on average compared to 18% for *Off*. We also see similar effects for the top-10% and top-20% videos (Figure 10).

6.3 Effect of Pre-Video Advertisements

To examine if pre-video advertisements affected user interest, we identify connections that last less than 30s for *Soc*, *Off*, and *Synd*. (*Off* and *Synd* used two advertisement segments per video which were either 5 seconds or 15 seconds. *Soc* used no embedded advertising.) Figure 21 shows the fraction of connections as a function of how long the users' stayed online. A very small time (less than 5 seconds) indicates that the user lost interest and quit viewing the video. Even though *Off* used embedded advertisements, these do not affect user interest significantly and the behavior is similar to *Soc* which had no advertisements. However, the advertisements have a non-trivial impact on *Synd*; 2.5% of users leave within 1 second and average user attrition rate is 1% (up from 0.5% for *Off* and *Soc*). *Off* and *Soc* both delivered popular videos in near real-time. User interest in live content was high enough that users were willing to tolerate the advertisements. *Synd* largely delivered non real-time snippets or highlights that had lower user interest; thus the advertisements had a higher impact on user attrition.

6.4 Implications

We see a clear impact of presentation model on access patterns, video popularity, and viewing persistence. For example, *Soc* with its unique user engaging method, brings much faster and much more concentrated access compared to *Off* and *Synd*. At the same time, *Soc* also brings flash crowds of much greater speed and magnitude. With the rapid growth of online video and social networks, we are already seeing the convergence these diverse media, and we expect this convergence to increase in the future. However, our understanding of how presentation models affect user behavior and thus affect a VoD system, is quite limited. We highlight this as a key problem area that deserves further study.

7. RELATED WORK

Live VoD is still an emerging area. We are not aware of any extensive measurements of live VoD systems at the scale we consider in this paper. However, there are several measurement studies of traditional VoD and P2P live streaming systems, which we discuss below.

VoD systems have attracted many research efforts. One of the first studies is by Griwodz et al., who use off-line video rental records to study video popularity [18]. They also present a request generation model to model time-of-day effects in user behavior. A number of other studies have been dedicated to the analysis of user behavior. Almeida et al. [9] and Costa et al. [16] focus on user behavior in the context of streaming video servers used by two educational systems (eTeach and BIBS). The work by Chesire et al. focuses on session duration, file popularity, and sharing patterns among clients [15]. Similarly, Chang et al. analyze

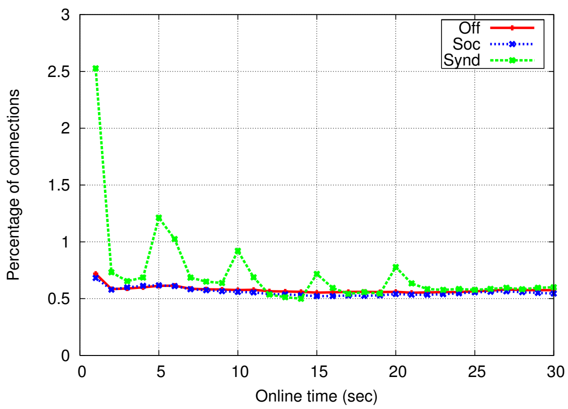


Figure 21: Connections between 1s and 30s

the user behavior in two media streaming servers run by Vanderbilt University media service, focusing primarily on file popularity, request inter-arrival time, and user viewing times [12].

Yu et al. give a comprehensive study of user behavior in a large scale VoD system deployed by China Telecom. They introduced a modified Poisson distribution for user arrival rates and also showed that the video popularity followed a Zipf distribution [28]. Huang et al. study long-term traces from the MSN VoD service and understand the implications of reducing server load via peer-assisted streaming [20]. They present an in-depth analysis of user behaviors and report results that users do not use interactivity features. They also present a theoretical framework to analyze the performance of peer-assisted VoD. Luo et al. analyze time-of-day effects, user interactivity, and popularity evolution using traces from a popular VoD system deployed by cctv.com, China's largest television station [23]. They use these results to evaluate the scalability of P2P approaches for VoD. Many other works focus on system design and user behaviors in P2P video streaming (e.g., [10, 13, 19]). There have been several recent efforts to provide DVD/VCR like functions (i.e., seek, rewind, pause) in P2P VoD services (e.g., [26, 27]). Such features might be more relevant in the context of traditional VoD services; our measurements indicate that users do not exercise these functions in a live VoD service.

With the advent of YouTube and other similar sites, user-generated content (UGC) has emerged as a primary driver for VoD services. Cha et al. provide an extensive measurement study of the user access patterns in YouTube [11]. They show that the popularity distribution exhibits power law like behaviors but with truncated tails. They also derive some implications for caching and system design from their analysis. An interesting observation in their work is that user preference seems relatively insensitive to a videos' age. In contrast, the live, event-driven nature of the Olympics dataset suggests significant churn in popular content. Cheng et al. [14] present measurements on how UGC systems differ from traditional VoD system, with a particular emphasis on the social networking aspects of the deployment, which they show has significant effects. Both these studies show that if a video did not attract enough requests when it first appears, then it is unlikely to attract too many requests in the future. This in contrast to the effects we saw from our case study of the flash crowd during the 100m finals. These measurements involve a user population that is much more temporally diverse than our dataset. Also, these studies do not have as much of a "white-box" view as we were able to provide in our measurements. Thus, they provide limited measurements on time-of-day effects, flash crowds, or how video access concentrations change over short time scales. Gill et al analyze YouTube usage at the University of Calgary network and observe time-of-day and day-of-week effects [17].

8. SUMMARY AND IMPLICATIONS

The 2008 Olympics saw a staggering scale of online live streaming never before seen on the Internet. The real-time, event-driven nature coupled with the varied presentation models imposes significant new demands on VoD systems.

We presented insights into this large-scale event through a unique dataset provided by ChinaCache, the largest CDN in China, which was responsible for serving video content for the three largest content providers in mainland China. This dataset provides us with a white-box view of the CDN servers serving the streaming media content to end-users. Using this dataset, we were able to understand: (1) how the live VoD workload differed from traditional VoD and UGC, (2) how user behaviors were affected by the real-time and event-driven nature of the event, (3) how the presentation models impacted video access and viewing patterns, and (4) case studies of how flash crowds manifest in such systems.

1. The real-time, event-driven nature results in patterns of user behavior and video access significantly different from traditional VoD systems. There are clear flash-crowd effects (caused by both expected and unanticipated events), the peak rush hour shows no time-of-day effects, video durations show more diversity, and video popularity changes much more dynamically.

2. User viewing times are largely independent of video durations and users prefer not to exercise DVD/VCR-like features such as seek or pause.

3. The presentation model affects access patterns significantly. A social networking site that actively engaged users and prompted users brought more concentrated access to popular videos and flash crowds of high intensity and magnitude.

4. Flash crowds can involve multiple related videos and earlier videos belonging to the same logical event are likely to get more concentrated access.

These observations suggest some guidelines for the design of future live VoD systems:

1. Simpler delivery systems (e.g., using HTTP) need not compromise user satisfaction. This may explain the rationale behind why the recent "smooth streaming" initiative uses HTTP delivery [8].

2. It might be more efficient to cache the first few minutes of many long videos instead of caching a few long videos in their entirety.

3. Content provider can leverage related videos to better provision for flash crowds and can defer releasing new content under overload.

4. Presentation models have significant implications for system design as they may bring more focused access to a small number of videos or result in flash crowds with a much smaller time-to-peak. We highlight this a key problem area that needs further study.

We do however caution system designers to take some of these observations and implications with a grain of salt. Olympics-style events present a worst-case scenario for system design – it is live, high-profile, event-driven, large-scale, and spans multiple days. At the same time, this singular nature also magnifies the importance of understanding the event in detail when the opportunity to study it presents itself. We were fortunate to have such an opportunity; we hope that our analysis provide such an understanding.

9. ACKNOWLEDGMENTS

This work is supported by the Project 60673184, 60873254 supported by NSFC, the Project 2007AA01Z419 supported by 863 Program, the Project 2008CB317101 supported by 973 Program, and the Tsinghua-ChinaCache CDN Research Institute Program. Hui Zhang and Vyas Sekar were supported in part by NSF award ANI-0331653. The authors would like to thank Huanying Zou, Shi Kang, Tongyu Zhan, Kunlong Wang, and Tongqing He for their excellent technical support.

10. REFERENCES

[1] Adobe Flash Media Server. `http://www.adobe.com/products/flashmediastreaming`.
[2] CCTV.com and Adobe Partner to Bring 2008 Beijing Olympics to Millions of Online Viewers in China. `http://www.adobe.com/aboutadobe/pressroom/pressreleases/pdfs/200808/080408AdobeCCTV.pdf`.
[3] ChinaCache CDN Principle. `http://en.chinacache.com/viewtechnique.asp?id=16`.
[4] ChinaCache Flash VoD. `http://en.chinacache.com/viewproduct.asp?id=344`.
[5] Data Center of the China Internet. `http://www.dcci.com.cn`.
[6] Facebook and CNN get together to stream inauguration. `http://blog.facebook.com/blog.php?post=48783697130`.
[7] Injury forces Liu Xiang to withdraw. `http://sports.espn.go.com/oly/summer08/trackandfield/news/story?id=3540221`.
[8] Smooth HD. `http://www.smoothhd.com/`.
[9] J. M. Almeida, J. Krueger, D. L. Eager, and M. K. Vernon. Analysis of Educational Media Server Workloads. In *Proceedings of Network and operating systems support for digital audio and video*, Jan. 2001.
[10] S. Annapureddy, S. Guha, C. Gkantsidis, D. Gunawardena, and P. R. Rodriguez. Is High-Quality VoD feasible using P2P Swarming? In *Proceedings of WWW*, May 2007.
[11] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon. I tube, You tube, Everybody tubes: Analyzing the World's Largest User Generated Content Video System. In *Proceedings of ACM SIGCOMM Conference on Internet Measurement*, Oct. 2007.
[12] B. Chang, L. Dai, Y. Cui, and Y. Xue. On Feasibility of P2P On-Demand Streaming via Empirical VoD User Behavior Analysis. In *Proceedings of Distributed Computing Systems Workshops, ICDCS'08*, June 2008.
[13] B. Cheng, X. Liu, Z. Zhang, and H. Jin. A Measurement Study of a Peer-to-Peer Video-on-Demand System. In *Proceedings of IPTPS*, Feb. 2007.
[14] X. Cheng, C. Dale, and J. Liu. Understanding the Characteristics of Internet Short Video Sharing: YouTube as a Case Study. To appear in *IEEE Transactions on Multimedia*, 2009.
[15] M. Chesire, A. Wolman, G. M. Voelker, and H. M. Levy. Measurement and Analysis of a Streaming Media Workload. In *Proceedings of USITS*, Mar. 2001.
[16] C. Costa, I. Cunha, A. Borges, C. Ramos, M. Rocha, J. Almeida, and B. Ribeiro-Neto. Analyzing Client Interactivity in Streaming Media. In *Proceedings of WWW*, May 2004.
[17] P. Gill, M. Arlitt, Z. Li, and A. Mahanti. YouTube Traffic Characterization: A View From the Edge. In *Proceedings of ACM SIGCOMM Conference on Internet Measurement*, Oct. 2007.
[18] C. Griwodz, M. Bar, and L. C. Wolf. Long-term Movie Popularity Models in Video-on-Demand Systems. In *Proceedings of ACM Multimedia*, Nov. 1997.
[19] Y. Guo, K. Suh, J. Kurose, and D. Towsley. P2Cast: Peer-to-peer Patching Scheme for VoD Service. In *Proceedings of WWW*, May 2003.
[20] C. Huang, J. Li, and K. Ross. Can Internet Video-on-Demand be Profitable. In *Proceedings of ACM SIGCOMM*, Aug. 2007.
[21] Y. Huang, T. Z. J. Fu, D. M. Chiu, J. C. S. Lui, and C. Huang. Challenges, Design and Analysis of a Large-scale P2P-VoD System. In *Proceedings of ACM SIGCOMM*, Aug. 2008.
[22] J. Jung, B. Krishnamurthy, and M. Rabinovich. Flash Crowds and Denial of Service Attacks: Characterization and Implications for CDNs and Web sites. In *Proceedings of WWW*, May 2002.
[23] J. G. Luo, Q. Zhang, Y. Tang, and S. Q. Yang. A Trace-Driven Approach to Evaluate the Scalability of P2P-Based Video-on-Demand Service. *IEEE Transactions on Parallel and Distributed Systems*, 20(1):59–70, Jan. 2009.
[24] V. N. Padmanabhan and K. Sripanidkulchai. The Case for Cooperative Networking. In *Proceedings of IPTPS*, Mar. 2002.
[25] K. Sripanidkulchai, B. Maggs, and H. Zhang. An Analysis of Live Streaming Workloads on the Internet. In *Proceedings of ACM SIGCOMM Conference on Internet Measurement*, Oct. 2004.
[26] A. Vlavianos, M. Iliofotou, and M. Faloutsos. Enhancing BitTorrent for Supporting Streaming Applications. In *Proceedings of IEEE Global Internet*, 2006.
[27] N. Vratonjic, P. Gupta, N. Knezevic, D. Kostic, and A. Rowstron. Enabling DVD-like features in P2P Video-on-Demand-Systems. In *Proceedings of Workshop on Peer-to-peer streaming and IP-TV*, Aug. 2007.
[28] H. Yu, D. Zheng, B. Y. Zhao, and W. Zheng. Understanding User Behavior in Large-Scale Video-on-Demand Systems. In *Proceedings of ACM Eurosys*, Apr. 2006.