Content Availability and Bundling in Swarming Systems

Daniel S. Menasche
University of Massachusetts

Antonio A. Aragao Rocha
Federal University of Rio de Janeiro

Bin Li
Tsinghua University

Don Towsley
University of Massachusetts

Arun Venkataramani
University of Massachusetts

ABSTRACT

BitTorrent, the immensely popular file swarming system, has a fundamental problem: availability. Although swarming scales well to tolerate flash crowds for popular content, it is less useful for unpopular or rare files as peers arriving after the initial rush find the content unavailable.

Our primary contribution is a model to quantify content availability in swarming systems. We use the model to analyze the availability and the performance implications of bundling, a strategy commonly adopted by many BitTorrent publishers today. We find that even a limited amount of bundling exponentially reduces content unavailability. Furthermore, for swarms with highly unavailable publishers, the availability gain of bundling can result in a net improvement in download time, i.e., peers obtain more content in less time. We empirically confirm the model's conclusions through experiments on PlanetLab using the mainline BitTorrent client.

1. INTRODUCTION

Despite the tremendous success of BitTorrent (estimated to account for 30–50% of all Internet traffic today), it suffers from a fundamental problem: availability. Although peer-to-peer content dissemination through *swarming* in BitTorrent scales impressively to tolerate massive flash crowds for popular content, swarming does little to disseminate unpopular content as their availability is limited by the presence of a seed or publisher. The extent of publisher unavailability is severe, e.g., our measurement study shows that 40% of the swarms have no publishers available more than 50% of the time.

University of Massachusetts at Amherst. July 2, 2009
.

To appreciate the availability problem, consider a swarm for an episode of a popular TV show. When a publisher first posts the episode, a flash crowd of peers joins the swarm to download it. The original publisher goes offline at some point, but peers may continue to obtain the content from other peers while the swarm is active. If a peer arrives after the initial popularity wave, when the population of the swarm has dwindled down to near-zero, though, it finds the content unavailable and must wait until a publisher reappears.

Our primary contribution is a mathematical model to study content availability in swarming systems such as BitTorrent. We use an $\mathrm{M}/\mathrm{G}/\infty$ queue to model the self-scaling property of BitTorrent swarms, i.e., more peers bring in more capacity to the system. The key insight is to model uninterrupted intervals during which the content is available as *busy periods* of that queue. The busy period increases exponentially with the arrival rate of peers and the time spent by peers in the swarm.

Our model also allows us to analyze the impact of *bundling*, a common strategy adopted by BitTorrent publishers wherein, instead of disseminating individual files via isolated swarms, a publisher packages a number of related files and disseminates it via a single larger swarm. To appreciate why bundling improves content availability, consider a bundle of K files. The popularity of the bundle is roughly K times the popularity of an individual file as a peer requesting any file requests the entire bundle. The size of the bundle is roughly K times the size of an individual file. Our model suggests that the busy period of the bundled swarm is a factor $e^{\Theta(K^2)}$ larger than that of an individual swarm. Indeed, if the busy period lasts until the publisher reappears, the content will be available throughout.

Surprisingly, in some cases, the improved availability can reduce the download time experienced by peers, i.e., peers download more content in less time. The *download time* of peers in the system consists of the *waiting time* spent while content is unavailable and the *service time* spent in actively downloading content. If the reduction in waiting time due to bundling is greater than the corresponding increase in service time, the download time decreases. We validate this conclusion in Section 4 through large-scale controlled experiments using the mainline BitTorrent client over Planetlab. Our controlled experiments also show that the conclusions of our model qualitatively hold even with realistic arrival patterns, peer upload capacities, and heterogeneous popularities.

In summary, we make the following contributions.

1) A large-scale *measurement study* showing that *(a)* content availability is a serious problem due to intermittently available publishers; and *(b)* bundling is widely prevalent and bundled content is more available.

2) A novel *analytic model* for content availability in swarming systems. Applying coverage processes to analyze peer-to-peer systems, our model is the first to account for waiting due to content unavailability.

3) A formal analysis of the performance and availability *implications of bundling*, a prevalent yet unexplored phenomenon in peer-to-peer systems, showing that it provides significant availability gains and reduces the download time for unpopular content with highly unavailable publishers. We empirically validate of the conclusions of the model based on controlled experiments with the mainline BitTorrent client on PlanetLab.

2. MEASURING CONTENT AVAILABILITY AND BUNDLING IN BITTORRENT

In this section, we present a large-scale measurement study of BitTorrent that shows that 1) content availability is a serious problem in BitTorrent today, and 2) bundling of content is widely prevalent and bundled content has higher availability. We begin with a brief overview of how swarming in BitTorrent works and why content becomes unavailable.

2.1 Why unavailability?

A swarm consists of a set of peers concurrently sharing (downloading or uploading) content (a file or a bundle of files) of common interest with the help of a coordinating tracker. Content is divided into blocks and peers obtain metadata about constituent blocks as well as identities of other peers in the swarm from the tracker. A peer exchanges blocks with other peers using a tit-for-tat incentive strategy until it completes its download. Peers that have not yet completed their download are called *leechers* while peers that possess all blocks in the content are called *seeds*.

Content is available if either at least one seed is present or sufficiently many active leechers are present so as to collectively make all constituent blocks available. Seeds may become unavailable in practice due to several reasons. Publishing sites serving a large number of files may take down seeds after the initial popularity wave subsides in order to reduce bandwidth costs. A seed may also be an average user publishing home-generated content that can not afford to stay online all the time. Seeds illegally uploading copyrighted material often disappear quickly for obvious reasons. Even for legitimate content, maintaining highly available seeds entails administrative effort and cost, which runs counter to the goals of content publishers that value BitTorrent as a cheap alternative to a client-server approach. [1]

Throughout this section, we measure content availability by equating it with seed availability. In the next section, we model content availability resulting both from seeds as well as from leechers alone.

2.2 Measuring unavailability

How available is content in BitTorrent swarms today? To answer this question, we conducted a seven-month long measurement study of BitTorrent swarms as follows. We developed and deployed BitTorrent monitoring agents at 300 nodes on Planetlab from August 3, 2008 to March 6, 2009. Once every hour, a host at the University of Massachusetts Amherst receives an RSS feed advertised by GoogleReader of recently created torrent URLs from Mininova (a large torrent hosting site), and sends each URL to a subset of the monitoring agents on Planetlab. The agents fetch the torrent metadata by joining the swarm and begin to monitor its peers. Our agents leverage the Peer Exchange (PEX) protocol extension, that enables it to discover new neighbors from other peers in addition to the tracker. To avoid copyright issues, our agents collect information only about the control plane without actually uploading or downloading content, which suffices for our purposes as we equate content availability with seed availability.

To distinguish seeds from leechers, our agents record the bitmaps received from connected peers. The bitmaps are part of the BitTorrent protocol and a peer uses them to convey the blocks it possesses to its neighbors. Each entry in the trace collected by the agents consists of a swarm identifier, a peer identifier (IP address and port number) and its bitmap recorded roughly periodically for each discovered peer in the swarm. Our traces consist of more than 14 million distinct IP addresses and 66K distinct swarms.

Figure 1 shows the distribution of seed availability for the monitored swarms. The solid curve shows the availability in the first month after the creation of the swarm, when we expect the content to be more popular. The extent of publisher unavailability is severe: less than 35% of the swarms had at least one seed available all the time. The availability of swarms over the entire duration of the measurement is even lower as shown by the dotted curve: almost 80% of the swarms are

[1]Henceforth, we use the terms *publishers* and *peers* interchangeably with *seeds* and *leechers* respectively.

unavailable 80% of the time.

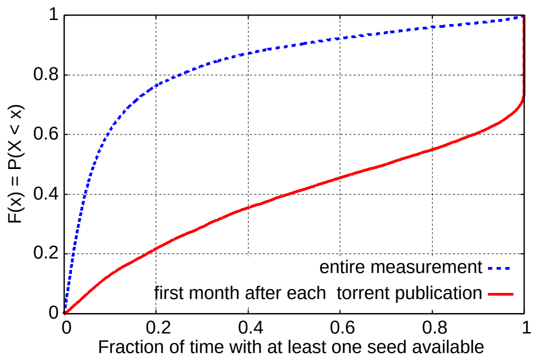

Figure 1: CDF of seed availability in 45,693 swarms each monitored for at least one month.

2.3 Content bundling

Bundling of content is a common practice in BitTorrent today. In this section, we study the extent of bundling and its impact on availability. The trace used in this section is a snapshot of BitTorrent swarms taken on May 6, 2009. For each of the 1,087,933 swarms in this snapshot, we record its content category (e.g., movies, TV, books etc.), names and sizes of constituent files, creation date, and instantaneous number of seeds and leechers.

2.3.1 Extent of bundling

We analyze the extent of bundling in three of the nine categories present in Mininova, namely, music, TV shows and books. These three categories together account for 45.98% of the swarms and 31.93% of the peers in the system. We chose these three categories because it is easier to automatically detect bundling by checking for the presence of multiple files with known extensions (e.g., `.mp3` for songs, `.mpg` for TV shows and `.pdf` for books). Detecting bundling is nontrivial in some categories, e.g., a DVD for a single movie is often organized as a collection of video files that are never distributed individually, making it difficult to check for the presence of multiple movies without manual inspection.

Among music swarms, albums are common. We classify a music swarm as a bundle if it has two or more files with common audio file extensions such as `.mp3`, `.mid` and `.wav`, which results in 193,491 of the 267,117 monitored swarms being classified as bundles.

Among TV show swarms, many bundles consist of sets of episodes in a season. We classify swarms that have two or more files with common video file extensions such as `.mpg` and `.avi` as bundles, which results in 25,990 of the 164,930 monitored swarms being classified as bundles.

Among book swarms, we observe that collections, i.e., torrents containing the keyword "collection" in their titles, usually consist of a bundle of contents connected by a broad theme, e.g., the "Ultimate Math Collection (1)" of size 5.81 GB has 642 books. We classified 841 of the 66,387 monitored swarms as collections. Classifying swarms that contain 2 or more files with common document file extensions such as `.pdf` and `.djvu` as bundles results in an additional 6,270 bundles.

2.3.2 Bundled content is more available

In this section, we present evidence suggesting that bundling is correlated with higher availability. We first consider book swarms. We find that 62% of all book swarms had no seed available on May 6, 2009, whereas that number drops to 36% if we consider only collections. Furthermore, the average number of downloads for a typical book swarm is 2,578, whereas for collections it is 4,216.

One reason for higher seed availability may be that content publishers are intrinsically more willing to support seeds for bundled content. The higher number of downloads for bundled content may be either because of higher demand for bundled content (as any peer seeking any of the constituent files may opt to download the bundle), or because of higher availability, or both. Higher seed availability in turn may in part be because of the increased number of downloads as some peers may choose to altruistically disseminate the content further. Although it is difficult to discern cause and effect in our measurement data, our analytic model in the next section quantifies how the higher demand and higher seed availability for bundled content *cause* improved content availability.

We next analyze our traces more closely for content that is available both in isolation and as part of a bigger bundle. We observe that among the unavailable collections, some of them were subsets of bigger collections, e.g., the 23 swarms consisting of collections of Garfield comics from 1978 to 2000 had no seeds. However, each of these collections can be found in a single super-collection aggregating all Garfield comics. The super-collection had seven seeds. After a manual inspection of all 841 book collections, we concluded that 210 had no seeds and were not subsets of other collections, which results in $210/841 = 25\%$ unavailability for content disseminated through collections (compared to 62% above for a typical swarm).

As another example, we consider swarms for the popular TV show "Friends". There were a total of 52 swarms associated with this show. Among them, 23 had one or more seeds available, and the remaining 29 had no seeds. The 23 available swarms consisted of 21 bundles (and 2 single episodes), whereas the 29 unavailable swarms consisted of only 7 bundles. These observations suggest a strong correlation between bundling and higher availability. The next section presents an

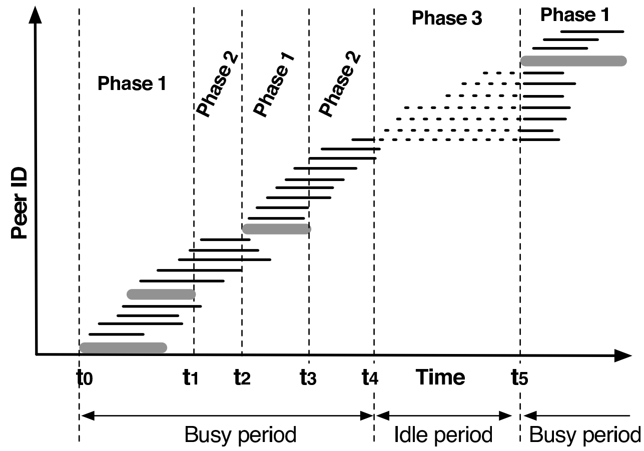

Figure 2: Illustration of busy and idle periods.

analytic model that quantifies the causal relationship between the two.

3. MODEL

In this section, we develop a model for content availability in BitTorrent. The key insight underlying the model is to view BitTorrent as a *coverage process* or equivalently an M/G/∞ queuing system. The model shows that 1) bundling improves availability, and 2) for swarms with highly unavailable publishers, the availability benefit of bundling more than offsets the increased time to actively download more content, resulting in a net decrease in user-perceived download times.

3.1 Model overview

Figure 2 illustrates how content availability in BitTorrent depends upon the arrivals and departures of publishers and peers. Each horizontal line segment represents the time interval during which a peer (represented using thin lines) or a publisher (represented using thick lines) stays online. A swarm is initiated by the arrival of a publisher, which also marks the start of the first *busy period.* The swarm's lifetime is divided into alternating busy and *idle periods.* Content is available during busy periods and unavailable during idle periods. If a publisher is always online, the first busy period lasts forever and content remains always available.

A busy period ends when the following two conditions are satisfied: 1) there are no publishers online, and 2) the *coverage*, i.e., the number of peers currently online, drops below a fixed small threshold (causing some blocks to become unavailable). For example, Figure 2 shows that after all publishers leave at time t_1, the busy period continues with the help of peers alone until a publisher reappears at time t_2. A busy period may alternate any number of times between a phase consisting of one or more publishers (Phase 1) and a phase consisting of peers alone (Phase 2). Peers arriving during either phase in a busy period will find the content available. At t_4, there are no publishers and the number of peers drops below the coverage threshold (assumed 3 in this example). This initiates an idle period that lasts until a publisher reappears at time t_5. Extant peers at the end of a busy period as well as peers arriving during the idle period find the content unavailable (represented by dotted lines). Because of idle waiting, these peers experience longer *download times* defined as the times since a peer arrives until it completes the download.

Our goal is to understand how content availability and the download times experienced by peers in a swarm depend upon 1) its popularity or the peer arrival rate λ; 2) the mean time s/μ that a peer takes during a busy period to actively download the content of size s at a rate equal to the effective average capacity μ of the swarm; and 3) the arrival rate r of publishers and the mean time u that a publisher stays online. We have implicitly assumed that u must be long enough for at least one copy of the file to be served in each busy period. For simplicity, we have assumed that peers are selfish and leave as soon as they complete their download; §3.3.4 extends the model to incorporate altruistic lingering.

To appreciate why bundling improves content availability, consider the special case of a highly unavailable publisher, i.e., its arrival rate r and mean residence time u are small. Then, the length of a busy period is determined primarily by peer arrivals and departures. Assuming Poisson peer arrivals and a coverage threshold of one, the length of a busy period can be shown to be $\frac{e^{\lambda s/\mu}-1}{\lambda}$. Bundling K files increases the peer arrival rate for the bundle to $K\lambda$ as each peer desiring any of the constituent files requests the entire bundle, and increases the time spent by each peer in the swarm to Ks/μ. As a result, the length of the busy period for the bundled swarm is $\frac{e^{K^2\lambda s/\mu}-1}{K\lambda}$, which translates to a reduction in unavailability by a factor $e^{\Theta(K^2)}$. For highly unavailable publishers, the availability gains of bundling can outweigh the cost of the increased time to download K times as much content resulting in a reduction in the download time, i.e., peers obtain more content in less time.

The rest of this section formalizes the above claims and derives closed-form expressions for the total download time experienced by peers with and without bundling. Unless otherwise stated, we assume that interarrival times of peers and publishers, residence time of publishers, and file download times are all exponentially distributed.

3.2 A simple model for content availability

We present a simple instance of the above model to analyze content availability and show that bundling improves availability. The model makes several simplifying assumptions (which we progressively relax in subse-

Variable	Description (units)
λ_k	peer arrival rate (1/s)
$\Lambda = \sum_{i=1}^{K} \lambda_k$	bundled peer arrival rate (1/s)
s_k	file size (bits)
$S = \sum_{i=1}^{K} s_k$	bundle size (bits)
μ	mean download rate of peers (bits/s)
r_k	arrival rate of publishers (1/s)
R	arrival rate of publishers for the bundle (1/s)
u_k	mean publisher residence time (s)
U	mean bundled publisher residence time (s)
Metric	**Description (units)**
P_k	unavailability
$\mathcal{P}$	unavailability of bundle
T_k	download time (s)
$\mathcal{T}$	bundle download time (s)

Table 1: Variables denoted by lower case characterize swarm $k \in \{1, 2, \cdots, K\}$ in isolation, while variables denoted by capital letters characterize the bundle of K files. Metrics for the swarms in isolation and for bundles are denoted by plain and stylized letters, respectively. Subscripts are dropped when homogeneous files are considered.

quent sections), but brings out the key insight underlying all of our results.

Assumptions: Content is available if and only if there is at least one publisher online. A peer arriving during an idle period finds the file unavailable and immediately leaves, i.e., it does not queue up until a publisher arrives. ⋄

Availability of an individual swarm.

In swarm k, let r_k and u_k be the arrival rate and residence time of publishers (refer to Table 1 for notation). Swarm k cycles through busy and idle periods, with average length $E[B_k]$ and $1/r_k$, respectively. The probability P_k that a peer arrives to swarm k to find the content unavailable is

$$P_k = \frac{1/r_k}{E[B_k] + 1/r_k}, \qquad k = 1, \ldots, K \tag{1}$$

and

$$E[B_k] = \frac{e^{r_k u_k} - 1}{r_k} \tag{2}$$

The above follows from classical results for the busy period of an M/G/∞ queue [19].

Availability of a bundled swarm.

Let R and U be the arrival rate and residence time of publishers for the bundle, respectively. The probability $\mathcal{P}$ that a peer arrives to find the content unavailable in the bundled swarm is

$$\mathcal{P} = \frac{1/R}{E[\mathcal{B}] + 1/R} \tag{3}$$

where the average length of a busy period for a bundle of K files is

$$E[\mathcal{B}] = \frac{e^{RU} - 1}{R} \tag{4}$$

Consider the special case when the publishers arrival rates are the same for all files, i.e., $r_k = r$ and their residence times are also the same, i.e., $u_k = u$ for all K files. If R and U scale as $R = Kr$ and $U = Ku$,

$$E[\mathcal{B}] = \frac{e^{K^2 RU} - 1}{KR} \tag{5}$$

$$\mathcal{P} = \frac{1/R}{(e^{K^2 RU} - 1)/(KR) + 1/R} \tag{6}$$

Note that $E[\mathcal{B}]$ is a factor $e^{\Theta(K^2)}$ larger than the corresponding value for an individual swarm. It can also be verified that $-\log P_k = \Theta(1)$ and $-\log \mathcal{P} = \Theta(K^2)$. Thus, bundling reduces the probability of not finding the content by a factor $e^{-\Theta(K^2)}$.

Availability with publishers and peers.

Assumptions: The busy period is defined w.r.t. a coverage threshold of one, i.e., a peer arriving during a busy period always finishes the download in that busy period and the last peer to finish ends the busy period. ⋄

Content may be available even if there are no publishers online. Let the aggregate arrival rate of peers and publishers to the individual swarm and to the bundle be $\lambda_k + r_k$ and $\Lambda + R$, respectively. We consider the special scenario in which publishers are willing to disseminate at most one copy of the file every time they return to a swarm. In this case, $u_k = s_k/\mu$ (assumption relaxed in the following section),

$$E[B_k] = \frac{e^{(r_k + \lambda_k)s_k/\mu} - 1}{r_k + \lambda_k}, \qquad k = 1, \ldots, K \tag{7}$$

and

$$E[\mathcal{B}] = \frac{e^{(R+\Lambda)S/\mu} - 1}{\Lambda + R} \tag{8}$$

If, for all K files, $\lambda_k = \lambda$ and $s_k = s$ then $\Lambda = K\lambda$ and $S = Ks$. The bundled busy period is $E[\mathcal{B}] = e^{\Theta(K^2)}$. Thus, bundling reduces the unavailability by $e^{-\Theta(K^2)}$ even if the bundled publisher arrival rate is equal to the publisher arrival rate of the individual swarms.

3.3 A model for content availability and download time

Next, we quantify content availability and the mean download time experienced by peers in the case where 1) peers may wait for content to become available, 2) the mean residence time of the publisher may differ from the service time of peers and 3) the coverage threshold may be greater than one. We begin by presenting the theoretical background required by our model.

Our results rely on those reported by Browne and Steele [2] on the busy period of an M/G/∞ queue where the customer initiating the busy period has an exceptional residence time. In what follows, we derive the expected busy period for the scenario of interest in this section.

Let customers arrive according to a Poisson process with rate β. The residence time of the customer initiating a busy period is draw from an exponential distribution with mean θ. The residence time of all other customers, X, takes the form of one of two exponentially distributed random variables, X_1 or X_2, with averages α_1 and α_2, respectively; $X = X_1$ with probability q_1 and $X = X_2$ with probability $q_2 = 1 - q_1$. The expected busy period is

$$E[B] = \theta + \sum_{i=1}^{\infty} \frac{\beta^i}{i!} \sum_{j=0}^{i} \binom{i}{j} \frac{q_1^j q_2^{i-j} \alpha_1^{1+j} \alpha_2^{1-j+i} \theta}{\alpha_1\alpha_2 + j\theta\alpha_2 + \theta\alpha_1 i - \theta\alpha_1 j} \quad (9)$$

The reader can find the derivation of equation (9) as well as the proofs of the results that follow in the Appendix. In the rest of this section, unless otherwise stated, we assume that all files have the same size and demand and that the publishers arrival rates and residence times are the same across all swarms. Assuming homogeneous swarms allows us to drop the subscripts of variables referring to individual swarms. In the Appendix we show that most of our results extend to the case where different swarms have different characteristics.

3.3.1 Availability with impatient peers

Assumptions: Publishers arrive to individual swarms at rate r and stay in the system for a mean time u. For the bundled swarm, publishers arrive with rate R and stay for a mean time U. Peers that arrive during an idle period leave immediately. ⋄

We are interested in determining the probability that a request leaves without being served. Denote this as P and $\mathcal{P}$ for the individual and bundled systems, respectively. Then

$$P = \frac{1/r}{E[B] + 1/r} \qquad \mathcal{P} = \frac{1/R}{E[\mathcal{B}] + 1/R} \quad (10)$$

The average busy period for each individual swarm, $E[B]$, is obtained from (9) by setting the parameters as follows: $\beta = \lambda + r$, $\theta = u$, $\alpha_1 = s/\mu$, $q_1 = \frac{\lambda}{\lambda+r}$, $\alpha_2 = u$.

For the bundled swarm, the aggregate peer arrival rate is $\Lambda = K\lambda$ and the size is $S = Ks$. The average busy period, $E[\mathcal{B}]$, is obtained from (9) as follows: $\beta = \Lambda + R$, $\theta = U$, $\alpha_1 = S/\mu$, $q_1 = \frac{\Lambda}{\Lambda+R}$, $\alpha_2 = U$.

The following lemma concerns the number of peers served in a busy period. Assuming that both the bundle publisher arrival rate, R, and publisher residence time, U, are independent of K, yields

LEMMA 3.1. *The number of peers served in a busy period, $E[\mathcal{N}]$, increase as $e^{\Theta(K^2)}$ by bundling K files.*

Note that this result is qualitatively similar to the case when publishers and peers stay online for the same mean time (Section 3.2).

We now consider the scenario where peers have skewed preferences. Given K contents, let p_k denote the probability that a request is for content k, $k = 1, \ldots, K$. Assume that $p_k = c/k^\delta$, $\delta > 0$ (Zipf's law). Letting Λ denote the aggregate peer arrival rate for all K swarms, the arrival rate for swarm k is $\lambda_k = p_k\Lambda$. Under the assumption that the time to download the bundle scales as K/μ, one can show that the lemma above still holds (details in the Appendix).

In the theorem below we relate the asymptotics of the busy period to the probability that a request is not served. Under the same assumptions of lemma 3.1 we have,

THEOREM 3.1. *(Availability theorem) Bundling K files together decreases unavailability by a factor $e^{\Theta(K^2)}$.*

In the result above the publisher arrival rate for the bundle, R, is assumed to be constant and independent of K. Nevertheless, in the Appendix we show that even if $R = \Omega(e^{-cK^2})$, $c > 0$, the availability of the bundle is still greater than the availability of the individual swarm by a factor $e^{\Theta(K^2)}$. When enough files are bundled, the long busy periods of the bundled swarm make it mostly self-sustaining. Peers can almost always download the content even in the absence of publishers.

3.3.2 Mean download time with patient peers

Assumptions: Peers that arrive during an idle period wait for a publisher to become available. The other assumptions are the same as in Section 3.3.1. ⋄

We wish to compare the download time of peers with and without bundling. To this aim, we first compute the average busy period length in an individual swarm, $E[B]$. When content is unavailable and a publisher arrives to start a busy period, the group of waiting peers immediately begins to be served. Neglecting the possible impact of this group of peers on the duration of the busy period, the average busy period $E[B]$ can be obtained from (9) by setting $\beta = \lambda + r$, $\alpha_1 = s/\mu$, $q_1 = \frac{\lambda}{\lambda+r}$, $\alpha_2 = \theta = u$. In the Appendix we also provide an expression for $E[B]$ accounting for the possible impact of the group of peers that begins to be served when the publisher arrives.

The average download time, $E[T]$, is given by

LEMMA 3.2. *The average download time of a file when peers are patient is*

$$E[T] = \frac{s}{\mu} + \frac{1}{r}P \quad (11)$$

where $P = \frac{1/r}{1/r+E[B]}$.

For the bundled swarm, the mean busy period length, $E[\mathcal{B}]$, can be obtained from (9) by setting $\beta = \Lambda + R$, $\alpha_1 = S/\mu$, $q_1 = \frac{\Lambda}{\Lambda+R}$, $\alpha_2 = \theta = U$. Once $E[\mathcal{B}]$ is obtained, the mean download time for the bundle, $E[\mathcal{T}]$, can be derived from (11) replacing s, r and $E[B]$ by their bundle counterparts S, R and $E[\mathcal{B}]$.

In the following theorem we relate the mean download times of bundles and individual swarms,

THEOREM 3.2. *(Download time theorem) Bundling K files can,*

(a) increase the mean download time by at most a factor K;

(b) decrease the mean download time of each file by a factor $\Theta(1/R)$ which grows unbounded as $R \to 0$.

Part (a) of the theorem above holds when service times dominate mean download times, in which case bundling decreases performance. Nevertheless, from part (b) of the theorem we conclude that if publishers become highly unavailable, peers may experience arbitrarily smaller download times when downloading bundles.

3.3.3 Threshold coverage

Assumptions: The assumptions are those described in §3.1. ⋄

If a peer leaves the system carrying the last copy of a chunk, content may become unavailable even if the number of peers online, i.e., the coverage, is greater than one. Our aim now is to determine the availability and the mean download time experienced by peers in the general case where content becomes unavailable when no publisher is online and the coverage reaches a threshold m. In Section 4.3.1 we validate the download time derived in this section against experiments.

Let $B(n, m)$ be the expected length of a *residual busy period* that begins with n leechers and ends as soon as the population size reaches m. The average busy period corresponds to $B(1, 0)$. $B(n, m)$ is given by

LEMMA 3.3. *For all n,*

$$B(n, 0) = \sum_{i=1}^{n} \frac{s}{i\mu} + \frac{s}{\mu} \sum_{i=1}^{\infty} \left(\frac{s\lambda}{\mu}\right)^i \frac{(n+i)! - n!i!}{i!(n+i)!i} \quad (12)$$

For $m < n$, $B(n, m)$ is obtained using the recursion $B(n, m) = B(n, 0) - B(m, 0)$.

We use lemma 3.3 to estimate the unavailability probability and the expected download time of peers in the scenario described in §3.1 and depicted in Figure 2. We assume that 1) the distribution of the residual busy period is concentrated around its mean and 2) publishers stay long enough in the system so that, when Phase 2 begins, the population of peers is in steady state. We denote the mean residual busy period starting when the system transitions to Phase 2 by $B(m)$,

$$B(m) = \sum_{i=0}^{\infty} \frac{e^{-\frac{\lambda s}{\mu}} \left(\frac{\lambda s}{\mu}\right)^i}{i!} B(i, m) \quad (13)$$

Noting that the number of times that the system cycles through Phases 1 and 2 before transitioning to Phase 3 is described by a geometric random variable with average $e^{rB(m)}$ yields

THEOREM 3.3. *For a threshold coverage of m, the probability that a request leaves without being served is*

$$P = \exp(-r(u + B(m))) \quad (14)$$

For patient peers, the expected download time is obtained by substituting (14) into (11). The corresponding expression for bundled swarms is obtained by replacing s, λ, r and u by their bundled counterparts, S, Λ, R and U. In particular, if $R = Kr$ and $U = Ku$ the availability and download time theorems still hold. In Section 4.3.1, we validate the mean download time estimated using (11) and (14) against experiments.

3.3.4 Altruistic lingering

Assumptions: Peers remain in the system for an average amount of time $1/\gamma$ after completing their downloads. The other assumptions are the same as in Section 3.3.2. ⋄

Peers may stay online as seeds after completing their downloads, either because they are altruistic or because publishers provide them incentives to do so. In the Appendix we show how to parameterize a general version of equation (9) to derive the availability probability and the mean download time of peers that stay online as seeds after completing their downloads. Furthermore, we also show that the availability and the download time theorems still hold.

To illustrate the consequences of peers staying longer in the system, consider two swarms with file sizes s_1 and s_2 and popularities λ_1 and λ_2. We wish to compare the performance of the individual swarms with that of a bundle with similar availability $[s_1\lambda_1/\mu + \lambda_1/\gamma = (\lambda_1 + \lambda_2)(s_1 + s_2)/\mu]$. The residence time for requestors of content 1 is equal to

$$\frac{s_1}{\mu} + \frac{1}{\gamma} = \frac{(\lambda_1 + \lambda_2)(s_1 + s_2)}{\mu\lambda_1} = \frac{s_1 + s_2}{\mu}\left(1 + \frac{\lambda_2}{\lambda_1}\right) \quad (15)$$

For the bundled swarm, the download time of peers is given by $(s_1 + s_2)/\mu$.

Assume swarm 1 is associated with a small and unpopular content while the swarm 2 content is large and

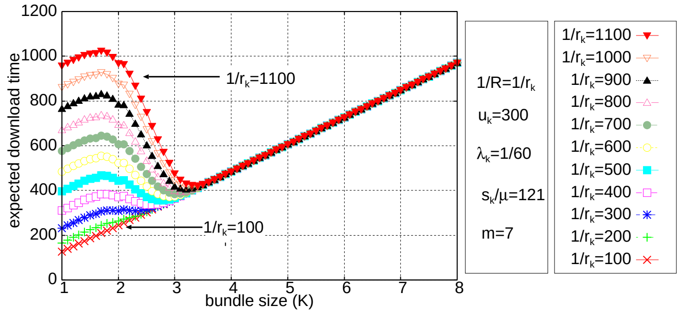

Figure 3: Bundles may reduce download time.

popular, $s_1 \ll s_2$, $\lambda_1 \ll 1 \ll \lambda_2$. Since content 1 is very unpopular (peer interarrival time very large), high availability depends on peers staying for a long time in the system after concluding their downloads (in equation (15), $1 + \lambda_2/\lambda_1 \to \infty$ as $\lambda_1 \to 0$). If swarm 1 is bundled with swarm 2, on the other hand, the overhead incurred by the peers only interested in content 2 is marginal (since $s_1 \ll s_2$) but the gains for peers interested in content 1 is remarkable, since requestors for content 1 experience the same availability and performance as those requesting file 2.

3.4 When can bundling reduce download time?

In this section we use the proposed model to illustrate when bundling reduces mean download time. We numerically evaluate equations (11) and (9) by setting the parameters as described in the legend of Figure 3. Figure 3 shows the expected download time as a function of the bundle size. For seven of the scenarios ($1/R \in [500 - 1100]$), increasing K to its optimal value, $K = 3$, leads to a decrease in the expected download time, while setting $K = 1$ is the best strategy for the remaining four. In each curve, as K increases the mean download time first increases, then decreases and finally increases again. The initial performance degradation occurs because small bundles may increase service times without sufficiently increasing the busy period. Figure 3 also shows that the benefits of bundling increase as the value of R decreases.

4. EXPERIMENTAL EVALUATION

In this section, we conduct controlled experiments using real BitTorrent clients to validate the two main conclusions of our model: 1) bundling improves availability, and 2) bundling can reduce download times when publishers are highly unavailable. We use an instrumented version of the mainline BitTorrent client [8] and experiment with private torrents deployed on Planetlab. Our experimental setup thus emulates realistic wide-area network conditions, client implementation artifacts, and the impact of realistic upload capacity distributions and arrival patterns that are difficult to capture in an analytic model.

4.1 Experimental setup

Our experiments were conducted using approximately 150 Planetlab hosts and two hosts at the University of Massachusetts Amherst one of which is designated as the controller of the experiment and another as a Bittorrent tracker. The controller causes peer arrivals, publisher arrivals, and publisher departures by dispatching via `ssh` a command to start or stop the BitTorrent client on a randomly chosen unused Planetlab host. At the end of the experiment, the controller collects the remote traces logged by the instrumented BitTorrent clients. Each client's trace logs the instantaneous download and upload rates every second as well as the fraction of the file downloaded up to that time.

Experimental parameters. Our experiments consist of torrents that publish either a single file of size $S = 4$ MB or a bundle of K files of aggregate size KS. The peer arrival rate for a bundle is assumed to be the sum of the arrival rates of its constituent files. The uplink capacity of each peer is $\mu = 33$ KBps ($\mu = 50$ KBps in §4.3). The publisher's upload capacity is 50KBps for individual as well as bundled torrents. There is only one publisher that alternates between being on and off. The peer arrival rate λ and on/off behavior of the publisher are varied according to the experimental goals as described below.

4.2 Bundling improves availability

Our model suggests that bundling reduces unavailability by increasing the length of busy periods and thereby reducing the reliance on a stable publisher. As an extreme case, we consider a publisher that initiates a swarm and then goes offline never to come back, and look at how long the swarm remains available after the publisher goes offline. We ensure that the publisher stays online long enough for at least one peer to fully download the file. Each peer leaves the system immediately after downloading the file.

We set λ=1/150 per second for each file and all other parameters to their default values, and study how the availability of the publisher-less swarm varies with the level of bundling K.

Figure 4 shows the number of peers served between 0 and 1500 seconds of the experiment for K=1, 2, 4, 6, 8 and 10. No peer completes its download in the first 300 seconds of the experiment: the publisher is either waiting for the first peer to arrive or is serving the first peer in each case. However, when the first peer completes its download and the publisher goes offline, the curves for $K = 1, 2, 4$ exhibit a very different trend compared to $K = 6, 8, 10$. For $K = 1, 2, 4$, only a small number of additional peers are able to complete their download before parts of the content start to become

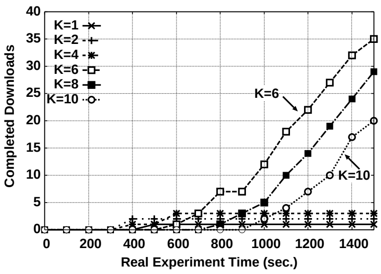

Figure 4: Availability of seedless swarms and the tradeoff in the choice of the bundle size.

unavailable. On the other hand, for $K = 6, 8, 10$, the number of completed downloads increases linearly, i.e., the swarm is self-sustaining even in the absence of a publisher.

In steady state, the length of time the swarm remains self-sustaining after the publisher goes offline is given by the mean residual busy period, $B(m)$. To compute $B(m)$ we use eq. (13) with $\mu = 33$KBps, $s = 4$MB and $\lambda = 1/150$ peers/s. A threshold coverage of $m = 9$ leads to the following values of $B(m)$ for $K = 1$ to 8, (0, 0, 47, 569, 2816, 8835, 256446, 75276). These values capture the fact that for $K \geq 5$ the swarms remained self-sustaining throughout our measurement.

Although the system goes from being unavailable to being available as K increases from 4 to 6, further increasing K only results in increased download times. The average download time of peers when $K = 10$ is roughly 66% higher than that for $K = 6$ (not shown in Figure 4). This suggests a delicate tradeoff in choosing K—it should be large enough to bridge gaps in publisher unavailability, but beyond that point bundling only increases download times. We study this tradeoff in more detail next.

4.3 Bundling can improve download time

In this section, we consider an intermittently available publisher with capacity 100KBps that alternately remains on and off for (exponentially distributed) mean times of 300s and 900s respectively. The arrival rate of peers for each file is $\lambda = 1/60$ per second and the capacity of each peer is μ=50 KBps. We study how the average download time of peers varies with the level of bundling.

Figures 5(a)–(c) show peer arrivals and departures over time. Each line segment starts at the instant that the peer arrives and terminates when the peer departs. For each value of K the experiment lasts for 10 runs of 1200s each. Figure 5(a) shows that for $K = 2$, many peers complete their download at roughly the same time. These flash departures indicate that the swarm is not self-sustaining. They happen because extant as well as newly arriving peers get stuck soon after the publisher goes off, and must wait until the publisher reappears and serves the missing blocks allowing them to complete their downloads On the other hand, setting K=3 (Figure 5(b)) reduces the likelihood of peers getting blocked, and setting $K = 4$ (Figure 5(c)) nearly eliminates blocking as the swarm sustains itself during periods of publisher unavailability.

Figure 6(a) shows the mean download time as function of K. For $K = 1$ and 2, the download time remains high as it is dominated by the time peers spend waiting for the publisher. The high variance is because of the variance in the downtime of the publisher. When $K = 3$, the mean download time reduces significantly, however the variance remains high as the download times are still partly determined by peers waiting for the publisher to reappear. The optimal bundle size is $K = 4$. The mean and the median download time as well as the variance are the lowest for this value of K as bundling eliminates gaps in publisher availability. For values of $K > 4$ the download time increases linearly with respect to K as the download time is dominated by the time to actively download increasingly bigger bundles. The variance continues to remain low as the swarm is increasingly self-sustaining with increasing K.

4.3.1 *Evaluation of the analytical model*

Next, we validate our analytical model (Section 3.3.3) against the experimental results above. We compute the download time using equation (11) and adapting (14) to account for the fact that there is only one publisher in the system to obtain

$$\mathcal{P} = \frac{\exp\left(-R\sum_{i=0}^{\infty}\frac{\exp(-\frac{K^2\lambda s}{\mu})(\frac{K^2\lambda s}{\mu})^i}{i!}B(i,m)\right)}{UR+1} \quad (16)$$

The derivation of the formula is in the Appendix. Setting $s/\mu = 80$s, $\lambda = 1/60$ peers/s, $1/r = 900$ arrivals/s, $u = 300$s and $m = 9$, our model predicts the results observed in Figure 6(a) pretty well. The model leads to an optimal bundle size of $K = 5$, whereas the optimal observed in the experiments was $K = 4$, and correctly captures the trend of the download time curve.

4.3.2 *Heterogeneous upload rates*

Next, we repeat the above experiment with heterogeneous peer upload capacities. The upload rate distribution was taken from the measured data used to generate Figure 1 in the BitTyrant study [12]. The average upload rate is 280KBps and the median is 50KBps. Using realistic peer upload capacities does not qualitatively change the behavior of the system (compare Figures 6(a) and Figures 6(b)). However, the optimal bundle size is now $K = 5$. This is consistent with the increase in the average upload capacity compared to the values obtained from the experiments with homo-

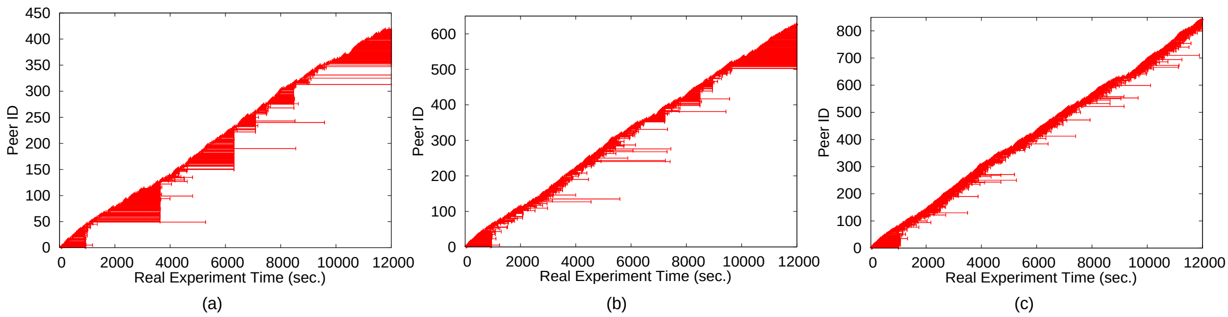

Figure 5: An intermittent publisher: (a) K=2; (b) K=3; (c) K=4. Each line starts when a peer arrives and ends when it leaves. As K increases, blocking probability decreases.

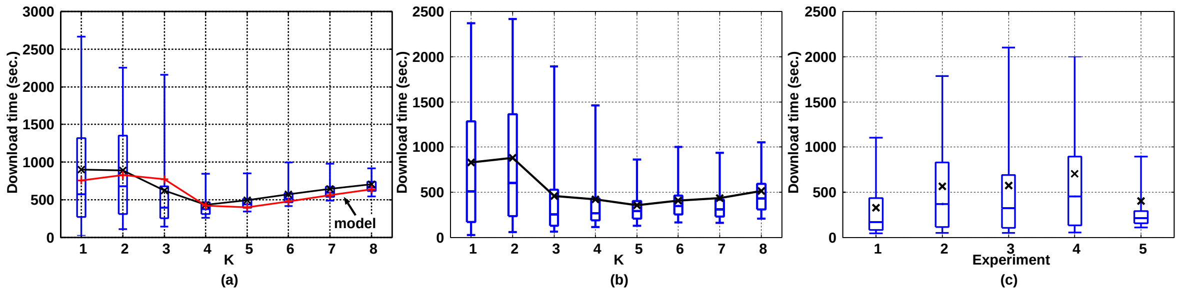

Figure 6: Download time versus bundling strategy. (a) exponential up and down times; (b) heterogeneous upload rates; (c) heterogeneous demand ($\lambda_i = \frac{1}{8i}$, $i = 1, \ldots, 4$), files bundled in experiment 5.

geneous capacities (μ=50 KBps). The higher upload capacity implies that a bigger bundle is needed to increase the length of its busy periods so as to make the swarm self-sustaining during periods of publisher unavailability—a conclusion that agrees with our model.

4.3.3 *Heterogeneous file popularites*

Next, we study the impact of bundling when different files have different popularities. We consider a bundle of $K = 4$ files. We assume that the popularities of the files inside the bundle are distributed as follows: $\lambda_1 = 1/8$, $\lambda_2 = 1/16$, $\lambda_3 = 1/24$ and $\lambda_4 = 1/32$. We run 5 experiments, the first four corresponding to swarms with individual files (experiments 1, 2, 3 and 4) and the last one to a bundle of all the files (experiment 5). In experiment i ($1 \leq i \leq 4$) we set λ_i as described above, and in experiment 5 we set $\lambda = \sum_{i=1}^{4} \lambda_i = 1/3.84$. All other parameters are set to their default values.

The mean download times are illustrated in Figure 6(c). The boxplots and lines show the distribution quartiles and 5th and 95th percentiles. For the individual files, as we move to the right in Figure 6(c) (i.e., as the popularity of the files decreases) the mean download time increases. When we consider a bundle of 4 files (experiment 5, extreme right in 6(c)) the mean download time is 405s. The mean download time of the bundle is larger than the download time of 329s experienced for file 1 in isolation but smaller than the download times for files 2, 3 and 4 in isolation. These results are explained as follows. File 1 is the most popular and stands little to gain in availability, so the cost of downloading more content outweighs the availability benefit of bundling. However, for the less popular files 2, 3 and 4, bundling reduces the download time by keeping the swarm self-sustaining during periods of publisher unavailability. In summary, if contents have different popularities, bundling may increase the download times of peers downloading the most popular contents but can benefit those downloading unpopular files. In this example, bundling slightly increases the download times of 48% of peers who download the most popular content but significantly benefits the majority of the population.

4.3.4 *Arrival patterns*

Our model as well as experiments so far assumed Poisson peer arrivals at a steady rate. To evaluate if our conclusions are sensitive to the Poisson assumption, we repeated experiments similar to those in Figure 6 using scaled versions of real arrival patterns observed in our measurement traces collected in §2. We found that using trace-driven arrivals does not qualitatively change our conclusions (refer to the Appendix for details).

However, we believe our model's conclusions may not

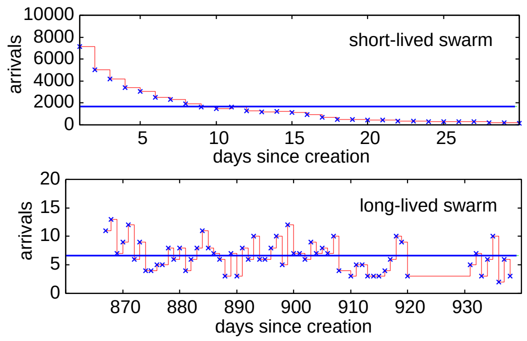

Figure 7: Typical peer arrival patterns of short-lived and long-lived swarms.

hold if the mean arrival rate is not steady for a long enough duration. In particular, our model will overestimate the length of the busy period and consequently availability if the arrival rate decreases significantly before the end of the busy period determined by the current arrival rate. Nevertheless, we found a significant number of swarms with relatively steady arrival rates in our measurement traces. For example, out of the 1,155 swarms associated with the TV show "Lost", 911 were published more than one month before we started our measurement. Figure 7(a) shows a typical new swarm in its first month and a typical old swarm after two years of its creation. The arrival rates of old swarms show much less variation compared to the arrival rates of new swarms. Our model can be used to predict the availability, download times, and the impact of bundling for such swarms.

5. RELATED WORK

A large body of prior work has investigated availability, performance and incentive issues in BitTorrent [3]. To our knowledge, this paper presents the first analytical model for content availability in BitTorrent-like swarming systems. We were also unable to find prior work studying the availability or performance implications of bundling in BitTorrent.

Ramachandran et al. [16] study the blocked leecher problem, where extant as well as arriving peers may have to wait for a long period of time for some blocks of the file that are no longer available. To address the problem, they propose BitStore, a token-based incentive architecture to obtain the missing blocks cached at other peers that had previously downloaded the file.

Neglia et al. [11] perform a large-scale measurement study to investigate availability in BitTorrent. They find that tracker availability is a serious enough problem that many torrents use replicated or DHT-based trackers for fault tolerance. Our focus is not on the availability of the tracker (or control plane), but on content availability (or data plane).

Both Susitaival et al. [18] and Wong et al. [21] relate the busy period of the $M/G/\infty$ queue to content availability in BitTorrent. Our model differs in two ways as it 1) quantifies content availability while accounting for publisher dynamics, and 2) quantifies the impact of bundling on availability and download time.

Qiu and Srikant [15] (building upon earlier work by Veciana et al. [20]) present a fluid model to analyze the download time performance of BitTorrent in steady-state. In contrast, our model that accounts for both performance and availability similar in spirit to *performability* [7]. A naive adaptation of the fluid model in [15] to bundles suggests strictly longer download times, whereas our model shows that bundling can improve download time by improving availability.

Many recent works have studied performance and fairness of a single swarm [8, 9, 10, 1, 4]. Collaboration across swarms was investigated by Guo et al. [6] suggesting many unexplored inter-torrent opportunities for block exchanges. Piatek et al. [13] suggest that propagating peer reputations limited to one hop can incent exchanges across swarms. Sirivianos et al. [17] propose an architecture where a commercial content provider provides "credits" to incent more cooperation between peers. Bundling is complementary to inter-swarm collaboration based on micropayment schemes for improving content availability. Micropayment schemes require a central bank to enable transactions and a tracking mechanism across swarms for peers to locate each other. In contrast, bundles are simple to set up and require no change to existing trackers or clients and is already in widespread use.

Economics of bundling.

Product bundling is a common commercial marketing strategy. The economics literature distinguishes between two forms of bundling [5]. In pure bundling or tying, a consumer can purchase the entire bundle or nothing at all. In mixed bundling, consumers have a choice to select parts of the package.

Both forms of bundling exist and have their pros and cons in BitTorrent's "bandwidth market" as well. Publishers can implement pure bundling by distributing bundled content as a zip archive. By forcing peers to download the whole bundle, pure bundling may make unpopular files more available, while subsidizing bandwidth costs for the publisher. However, it can delay those seeking exclusively popular files by forcing them to download content they do not want.

Mixed bundling is more common and can also improve availability. Publishers typically bundle files according to user interests, thus bundling can serve as a mutually beneficial recommendation system. A user seeking one episode of a TV show may decide to fetch the entire season for possible future viewing. A publisher might recommend a movie as part of a bundle to a user who may preview it and choose to pay for it after all [14, 21]. Even a small fraction of users opting to download more content than they strictly sought can significantly improve availability. Both mixed and pure bundling in BitTorrent have a beneficial side-effect: they replicate unpopular or rare content implicitly increasing their durability in the long run, i.e., it reduces the likelihood of rare content being lost permanently.

Bundling may increase the traffic in the network. Future work consists of studying the implications of bundling for the ISPs and its impact on content locality.

6. CONCLUSIONS

Peer-to-peer swarming in BitTorrent scales impressively to tolerate massive flash crowds, but falls short on availability. Although it is fashionable to observe that BitTorrent accounts for up to half of all Internet traffic today, it is less well known that half of the swarms are unavailable half of the time—an observation that does not bode well for the increasing commercial interest in integrating swarming with server-based content dissemination. Our work is a first step towards developing a foundational understanding of content availability in swarming systems.

By viewing BitTorrent as a queueing system, we were able to model content availability. The model suggests two important implications for bundling of content, a common practice among swarm publishers today. First, bundling improves content availability. Second, when the publisher is highly unavailable, bundling reduces the download time experienced by peers to obtain unpopular content. The latter implication is particularly intriguing as peers take less time to download more content. Although the model makes several simplifying assumptions, we were able to empirically validate its conclusions through large-scale controlled experiments with the mainline BitTorrent client over Planetlab.

Acknowledgement: We thank Giovanni Neglia and Shlomo Zilberstein for fruitful discussion on the economic implications of bundling. This work was supported in part by the NSF under award numbers CNS-0519922 and CNS-0721779. Research of DSM also funded in part by a scholarship from CAPES/Fulbright (Brazil).

7. REFERENCES

[1] Bharambe, A. R., Herley, C., and Padmanabhan, V. N. Some observations on bittorrent performance. In *SIGMETRICS* (2005).

[2] Browne, S., and J.M.Steele. Transient behavior of coverage processes with applications to the infinite server queue. *J. Appl. Prob. 30 (3)* (1993), 589–601.

[3] Cohen, B. Incentives build robustness in bittorrent. In *Workshop on Economics of Peer-to-Peer Systems* (2003).

[4] Fan, B., Chiu, D., and Lui, J. Stochastic analysis and file availability enhacement for bittorrent-like sharing systems. In *IWQoS* (2006).

[5] Fuerderer, R., Herrmann, A., and Wuebker, G. *Optimal Bundling*. Springer, 1999.

[6] Guo, L., Chen, S., Xiao, Z., Tan, E., Ding, X., and Zhang, X. A performance study of bittorrent-like peer-to-peer systems. *JSAC 25(1)* (2007), 155–169.

[7] Haverkort, B. R., Marie, R., and Rubino, G. *Performability Modelling*. Wiley, 2001.

[8] Legout, A., Liogkas, N., and Kohler, E. Rarest first and choke algorithms are enough. In *IMC* (2006).

[9] Legout, A., Liogkas, N., Kohler, E., and Zhang, L. Clustering and sharing incentives in bittorrent systems. In *SIGMETRICS* (2007).

[10] Massoulie, L., and Vojnovic, M. Coupon replication systems. In *SIGMETRICS* (2006).

[11] Neglia, G., Reina, G., Zhang, H., Towsley, D., and Arun Venkataramani, J. D. Availability in bittorrent systems. In *INFOCOM* (2007).

[12] Piatek, M., Isdal, T., Anderson, T., Krishnamurthy, A., and Venkataramani, A. Do incentives build robustness in bittorrent? In *NSDI* (2007).

[13] Piatek, M., Isdal, T., Krishnamurthy, A., and Anderson, T. One hop reputations for peer to peer file sharing workloads. In *NSDI* (2008).

[14] Pucha, H., Roy, S., and Hu, Y. C. Take one get one free; leveraging p2p networks for content promotion. In *Global Internet Symposium* (2006).

[15] Qiu, D., and Srikant, R. Modeling and performance analysis of bittorrent-like peer-to-peer networks. In *SIGCOMM* (2004).

[16] Ramachandran, A., das Sarma, A., and Feamster, N. Bitstore: An incentive compatible solution for blocked downloads in bittorrent. In *NetEcon* (2007).

[17] Sirivianos, M., Park, J. H., Chen, R., and Yang, X. Free-riding in bittorrent networks with the large view exploit. In *IPTPS* (2007).

[18] Susitaival, R., Aalto, S., and Virtamo, J. *Lecture Notes in Computer Science*. 2006, ch. Analyzing the dynamics and resource usage of P2P file sharing by a spatio-temporal model.

[19] Takacs, L. *Introduction to the Theory of Queues*. Oxford University Press, 1962.

[20] Veciana, G., and Yang, X. Fairness, incentives and performance in peer to peer networks.

[21] Wong, S., Altman, E., and Ibrahim, M. P2P networks: interplay between legislation and information technology. In *INRIA rr6889* (2009).

APPENDIX

Background

Our results rely on those reported by Browne and Steele [2] on the busy period of an M/G/∞ queue where the customer initiating the busy period has an exceptional residence time.

Let customers arrive according to a Poisson process with rate β. If we allow customers initiating a busy period to draw their residence times from a distribution $H(\cdot)$ with Laplace transform $h(\cdot)$ and mean θ while all other customers draw their residence times from a distribution $G(\cdot)$, the expected busy period length is given by

$$E[B] = \theta + \sum_{i=1}^{\infty} \frac{\beta^i}{i!} \int_0^{\infty} (1-H(x)) \Big[\int_x^{\infty} \Big(1-G(u)\Big) du \Big]^i dx \tag{17}$$

When $G(x) = 1 - e^{-x/\alpha}$, i.e., all customers except the first draw their service times from an exponential distribution, the equation above reduces to

$$E[B] = \theta + \sum_{i=1}^{\infty} \frac{(\beta\alpha)^i \alpha [1 - h(i/\alpha)]}{i! i} \tag{18}$$

If the customer initiating a busy period also draws its service time from an exponential distribution,

$$E[B] = \theta + \alpha\theta \sum_{i=1}^{\infty} \frac{(\beta\alpha)^i}{i!(\alpha + i\theta)} \tag{19}$$

Finally, if $\theta = \alpha$,

$$E[B] = (e^{\beta\alpha} - 1)/\beta \tag{20}$$

Proofs

Throughout the proofs, let $\lambda_S = \max\{\lambda_k\}$, $\lambda_I = \min\{\lambda_k\}$, $s_S = \max\{s_k\}$, $s_I = \min\{s_k\}$.

Derivation of equation (9)

PROOF. We use equation (17) to obtain (9). Let the download time of customers that arrive during the busy period be given by

$$X = \begin{cases} X_1 & \text{with probability } q_1 \\ X_2 & \text{with probability } q_2 = 1 - q_1 \end{cases}$$

where $E[X_i] = \alpha_i$. Then,

$$G(u) = 1 - q_1 e^{-\frac{1}{\alpha_1} u} - q_2 e^{-\frac{1}{\alpha_2} u} \tag{21}$$

and

$$E[B] = \theta + \sum_{i=1}^{\infty} \frac{\beta^i}{i!} \int_0^{\infty} I(z,i) dz \tag{22}$$

where

$$I(z,i) = (1-H(z)) \Big[\int_z^{\infty} \Big(q_1 e^{-\frac{u}{\alpha_1}} + q_2 e^{-\frac{u}{\alpha_2}} \Big) du \Big]^i \tag{23}$$

$$I(z,i) = (1 - H(z)) \Big[q_1 \frac{e^{-\frac{1}{\alpha_1} z}}{\frac{1}{\alpha_1}} + q_2 \frac{e^{-\frac{1}{\alpha_2} z}}{\frac{1}{\alpha_2}} \Big]^i \tag{24}$$

$$I(z,i) = (1-H(z)) \sum_{j=0}^{i} \binom{i}{j} \Big[\Big(q_1 \frac{e^{-\frac{1}{\alpha_1} z}}{\frac{1}{\alpha_1}} \Big)^j \Big(q_2 \frac{e^{-\frac{1}{\alpha_2} z}}{\frac{1}{\alpha_2}} \Big)^{i-j} \Big] \tag{25}$$

Therefore, $\int_0^{\infty} I(z,i) dz$ equals

$$\sum_{j=0}^{i} \binom{i}{j} \frac{q_1^j}{\frac{1}{\alpha_1}^j} \frac{q_2^{i-j}}{\frac{1}{\alpha_2}^{i-j}} \Big[\frac{1}{\frac{j}{\alpha_1} + \frac{i-j}{\alpha_2}} - \int_0^{\infty} H(z) e^{-(\frac{j}{\alpha_1} + \frac{i-j}{\alpha_2}) z} dz \Big] \tag{26}$$

Using integration by parts to solve (26) and substituting the result into (22) leads to

$$E[B] = \theta + \sum \frac{\beta^i}{i!} \sum_{j=0}^{i} \binom{i}{j} \frac{q_1^j}{\frac{1}{\alpha_1}^j} \frac{q_2^{i-j}}{\frac{1}{\alpha_2}^{i-j}} \Big[\frac{1 - h(\frac{j}{\alpha_1} + \frac{i-j}{\alpha_2})}{\frac{j}{\alpha_1} + \frac{i-j}{\alpha_2}} \Big] \tag{27}$$

and if $h(s) = \theta^{-1}/(\theta^{-1} + s)$,

$$E[B] = \theta + \sum_{i=1}^{\infty} \frac{\beta^i}{i!} \sum_{j=0}^{i} \binom{i}{j} \frac{q_1^j q_2^{i-j} \alpha_1^{1+j} \alpha_2^{1-j+i} \theta}{\alpha_1 \alpha_2 + j\theta\alpha_2 + \theta\alpha_1 i - \theta\alpha_1 j} \tag{28}$$

□

Proof of lemma 3.1

PROOF. Since all the terms in B_k, $k = 1, \ldots, K$, are lower bounded and upper bounded by terms that do not depend on K, one can show that $E[B_k]$ is bounded hence $B_k = \Theta(1)$.

To show that $\log \mathcal{B} = \Theta(K^2)$ we consider a special process where the following conditions hold,

- during the busy period, customers arrive with rate $\beta = R + \sum_{k=1}^{K} \lambda_k$,
- the residence times of all customers, including the first, arriving in a busy period are drawn from an exponentially distributed random variable with mean α,
- α and β are upper bounded by

$$\alpha \leq \frac{K s_S}{\mu} \tag{29}$$
$$\beta \leq K\lambda_S + R \tag{30}$$

- α and β are lower bounded by

$$\alpha \geq \frac{K s_I}{\mu} \tag{31}$$
$$\beta \geq K\lambda_I \tag{32}$$

The average busy period, $E[\mathcal{B}^\star]$, of the special process is given by (20)

$$E[\mathcal{B}^\star] = (e^{\beta\alpha} - 1)/\beta \tag{33}$$

First, we show that $\log E[\mathcal{B}^\star] = O(K^2)$.

$$E[\mathcal{B}^\star] \leq \frac{e^{K^2 s_S \lambda_S/\mu + o(K^2)} - 1}{K\lambda_I} \tag{34}$$

$$\lim_{K\to\infty} \frac{\log E[\mathcal{B}^\star]}{K^2} < \infty \tag{35}$$

Next, we show that $\log E[\mathcal{B}^\star] = \Omega(K^2)$.

$$E[\mathcal{B}^\star] \geq \frac{e^{K^2 s_I \lambda_I/\mu} - 1}{K\lambda_S + R} \tag{36}$$

$$\lim_{K\to\infty} \frac{K^2}{\log E[\mathcal{B}^\star]} < \infty \tag{37}$$

Therefore,

$$\log E[\mathcal{B}^\star] = \Theta(K^2) \tag{38}$$

To extend the result above to the parameterization of $E[\mathcal{B}]$ made in Section 3.3.1 we proceed as follows.

For the upper bound, $\log E[\mathcal{B}] = O(K^2)$, consider a modified process in which the residence times of all customers arriving during a busy period are drawn from an exponential random variable with mean $\alpha = \max(U, \frac{Ks_S}{\mu})$. Denote the busy period of the modified process by $\widehat{B}$. For K sufficiently large, $\alpha = \frac{Ks_S}{\mu}$. Noting that conditions (29)-(32) hold it follows that $\log E[\widehat{B}] = O(K^2)$. Since $E[\widehat{B}] \geq E[\mathcal{B}]$, $\log E[\mathcal{B}] = O(K^2)$.

For the lower bound, $\log E[\mathcal{B}] = \Omega(K^2)$, consider a modified process in which the residence times of all customers arriving during a busy period are drawn from an exponential random variable with mean $\alpha = \min(U, \frac{Ks_I}{\mu})$. Denote the busy period of the modified process by $\widetilde{B}$. For K sufficiently large, $\alpha = \frac{Ks_I}{\mu}$. Noting that conditions (29)-(32) hold it follows that $\log E[\widetilde{B}] = \Omega(K^2)$. Since $E[\widetilde{B}] \leq E[\mathcal{B}]$, $\log E[\mathcal{B}] = \Omega(K^2)$.

Finally, one can show that the results above extend to $E[\mathcal{N}] = E[\Lambda\mathcal{B}]$,

$$\log E[\mathcal{N}] = \Theta(K^2) \tag{39}$$

□

Proof of remark following lemma 3.1

In the following proof, let $r_k = r$ and $\lambda_k = \Lambda p_k$ for all $k = 1,\ldots,K$. Λ scales as $K\lambda$ and service times scale as K/μ. Without loss of generality, $s_k = 1$ for all $k = 1,\ldots,K$.

Proof. We begin considering a process similar to the special one described in lemma 3.1 to characterize the bundled swarm,

- during the busy period, customers arrive with rate $\beta = R + K\lambda$,
- the residence times of all customers arriving in a busy period are drawn from an exponentially distributed random variable with mean $\alpha = K/\mu$.

The busy period of this process is denoted by $\mathcal{B}^\star$. For the individual swarm k,

- during the busy period, customers arrive with rate $\beta = r + \lambda_k = r + K\lambda p_k$,
- the residence times of all customers arriving in a busy period are drawn from an exponentially distributed random variable with mean α,
- α satisfies $\min(1/\mu, u) \leq \alpha \leq \max(1/\mu, u)$.

The busy period of this process is denoted by $\mathcal{B}_k^\star$.

The probability that a request is for content k is p_k,

$$p_k = \frac{c}{k^\alpha}, \quad \alpha > 0, \quad k = 1,\ldots,K \tag{40}$$

where

$$c^{-1} = \sum_{k=1}^{K} p_k = \sum_{k=1}^{K} 1/k^\alpha \tag{41}$$

which is known to diverge if $\alpha \leq 1$,

$$\sum_{k=1}^{K} 1/k^\alpha = O(\log K), \qquad \alpha \leq 1 \tag{42}$$

and converge when $\alpha > 1$. In particular, $p_1 \to 0$ if $\alpha \leq 1$ and $p_1 \to c > 0$ if $\alpha > 1$.

Let $\eta_k^\star$ be the ratio of the mean busy period duration of the bundled swarm over the mean busy period duration of individual swarm k,

$$\eta_k^\star = \frac{E[\mathcal{B}^\star]}{E[B_k^\star]}, \qquad k = 1,\ldots,K \tag{43}$$

and let $\eta^\star$ be the ratio of the mean busy period duration of the bundled swarm over the mean busy period across all individual swarms,

$$\eta^\star = \frac{E[\mathcal{B}^\star]}{\sum_{k=1}^{K} p_k E[B_k^\star]} \tag{44}$$

Note that

$$\eta_k^\star = \frac{(e^{K^2\lambda/\mu + o(K^2)} - 1)/(R + K\lambda)}{(e^{K\lambda p_k/\mu + o(K^2)} - 1)/(r + K\lambda p_k)} \tag{45}$$

yields, after some algebraic manipulation, $\log \eta_k^\star = \Theta(K^2)$.

Similarly,

$$\eta^\star = \frac{(e^{K^2\lambda/\mu + o(K^2)} - 1)/(R + K\lambda)}{\sum_{k=1}^{K} p_k (e^{K\lambda p_k/\mu + o(K^2)} - 1)/(r + K\lambda p_k)} \tag{46}$$

yields $\log \eta^\star = \Theta(K^2)$. The later follows from

$$\begin{aligned} \log \eta^\star \;&=\; \log(e^{K^2\lambda/\mu + o(K^2)} - 1) - \log c - \\ & \log\Big(\sum_{k=1}^{K} \frac{e^{K\lambda p_k/\mu + o(K^2)} - 1}{(r + K\lambda p_k)k^\alpha}\Big) + o(K^2) \end{aligned} \tag{47}$$

Next, we show that all the terms in the formula above, except the first, are $o(K^2)$. First, we show that $-\log c = o(K^2)$. Let

$$c_\infty = \lim_{K\to\infty} -\log c = \lim_{K\to\infty} -\log\Big((\sum_{k=1}^{K} 1/k^\alpha)^{-1}\Big) \quad (48)$$

c_∞ is $O(\log K)$ if $\alpha \le 1$ and $c_\infty < \infty$ otherwise. For all α,

$$-\log c = -\log\Big((\sum_{k=1}^{K} 1/k^\alpha)^{-1}\Big) = o(K^2) \qquad (49)$$

Then, we show that

$$\log\Big(\sum_{k=1}^{K} \frac{e^{K\lambda p_k/\mu + o(K^2)} - 1}{(r+K\lambda p_k)k^\alpha}\Big) = o(K^2) \qquad (50)$$

Equation (50) follows from

$$\frac{e^{o(K^2)}-1}{(r+K\lambda)k^\alpha} \le \frac{e^{K\lambda p_k/\mu + o(K^2)} - 1}{(r+K\lambda p_k)k^\alpha} \le \frac{e^{K\lambda/\mu + o(K^2)}}{rk^\alpha} \qquad (51)$$

$$\log\Big(\frac{e^{o(K^2)}-1}{(r+K\lambda)}\sum_{k=1}^{K}\frac{1}{k^\alpha}\Big) \le \log\Big(\sum_{k=1}^{K}\frac{e^{K\lambda p_k/\mu+o(K^2)}-1}{(r+K\lambda p_k)k^\alpha}\Big)$$

$$\le \log\Big(\frac{e^{K\lambda/\mu+o(K^2)}}{r}\sum_{k=1}^{K}\frac{1}{k^\alpha}\Big) \qquad (52)$$

Equations (42) and (52) yield (50). Putting together (47), (49) and (50) we find that $\log\eta^\star = \Theta(K^2)$.

To extend the result above to the parameterization of $E[\mathcal{B}]$ made in Section 3.3.1 we proceed as in the conclusion of lemma 3.1. To illustrate, let us prove that

$$\log \eta_k = \log \frac{E[\mathcal{B}]}{E[B_k]} = O(K^2) \qquad (53)$$

Consider a process in which the residence times of all customers arriving during a busy period are drawn from an exponential random variable with mean $\alpha = \max(U, \frac{K}{\mu})$. Denote the busy period of the process by $\widehat{B}$. For K sufficiently large, $\alpha = \frac{K}{\mu}$. Consider a second process in which the residence times of all customers arriving during a busy period are drawn from an exponential random variable with mean $\alpha' = \min(u, \frac{1}{\mu})$. Denote the busy period of this second process by $\widehat{\widehat{B}}$. Equation (45) yields $\log\frac{E[\widehat{B}]}{E[\widehat{\widehat{B}}]} = \Theta(K^2)$. Since $E[\widehat{B}] \ge E[\mathcal{B}]$ and $E[\widehat{\widehat{B}}] \le E[B_k]$, $\log\eta_k = O(K^2)$. $\square$

Proof of theorem 3.1

PROOF. Since all the terms in P_k, $k = 1, \ldots, K$, are lower bounded and upper bounded by terms that do not depend on K, one can show that $E[B_k]$ is bounded hence $P_k = \Theta(1)$.

We rewrite $-\log\mathcal{P}$ as

$$\begin{aligned} -\log\mathcal{P} &= -\log\frac{1/R}{E[\mathcal{B}]+1/R} \\ &= -\log(1/R) + \log(E[\mathcal{B}]+1/R) \\ &= -\log(1/R) + \log(e^{\Theta(K^2)}+1/R) \quad (54)\end{aligned}$$

where the last equality follows from lemma 3.1.

We now show that $-\log\mathcal{P} = \Theta(K^2)$. First, we prove that $-\log\mathcal{P} = O(K^2)$,

$$\lim_{K\to\infty}\frac{-\log\mathcal{P}}{K^2} = \kappa_1 + \lim_{K\to\infty}\frac{\log(e^{\Theta(K^2)}+1/R)}{K^2} < \infty \qquad (55)$$

Then, we show that $-\log\mathcal{P} = \Omega(K^2)$,

$$\lim_{K\to\infty}\frac{K^2}{-\log\mathcal{P}} = \lim_{K\to\infty}\frac{1}{\kappa_2 + \frac{\log(e^{\Theta(K^2)}+1/R)}{K^2}} < \infty \quad (56)$$

from which we conclude that $-\log\mathcal{P} = \Theta(K^2)$. $\square$

Proof of remark following theorem 3.1

PROOF. Assume that R decreases as K increases. If $R = \Omega(e^{-cK^2})$ then $1/R = O(e^{cK^2})$ and the proof of theorem 3.1 holds.

If R increases as K increases, conditions (29)- (30) in lemma 3.1, used to prove that $\log E[\mathcal{B}] = O(K^2)$, may not hold. In particular,

$$\alpha\beta = K^2\frac{s_S\lambda_S}{\mu} + o(K^2) \qquad (57)$$

may not hold. Nevertheless, one can still show that $\log E[\mathcal{B}] = \Omega(K^2)$ and $\mathcal{P} = e^{-\Omega(K^2)}$. $\square$

Proof of lemma 3.2

Consider the group of peers that wait for a publisher to arrive and immediately begins to be served when the busy period starts. In the following proof, we account for the possible impact of this group of peers on the duration of the busy period.

PROOF. Let Q denote the queueing time of a request. The average queueing time conditioned on the event that a request queues is

$$E[Q|Q>0] = 1/r \qquad (58)$$

Let T denote the time required to download content. It's average, $E[T]$, is

$$\begin{aligned} E[T] &= P(Q=0)\frac{s}{\mu} + P(Q>0)\Big(\frac{s}{\mu} + E[Q|Q>0]\Big) \\ &= \frac{E[B]}{\frac{1}{r}+E[B]}\frac{s}{\mu} + \frac{1/r}{\frac{1}{r}+E[B]}\Big(\frac{s}{\mu}+E[Q|Q>0]\Big) \\ &= \frac{s}{\mu} + \frac{1/r}{\frac{1}{r}+E[B]}\frac{1}{r}\end{aligned}$$

where $E[B]$ is the expected length of a busy period of an M/G/∞ queue with special distribution for the first customer, obtained using equation (18), as detailed next.

When a publisher arrives to start a busy period a group of peers that were waiting for a publisher immediately begins to be served. This group is modeled using a virtual customer, with residence time Y. Y is the maximum of all the download times of the queued peers that enter simultaneously into service and the residence time of the publisher,

$$Y = \max\{X_1, \ldots, X_L, X_{L+1}\}$$

where $X_1, \ldots, X_L$ are exponential random variables with mean s/μ and X_{L+1} is exponential with mean u. L is a geometric random variable with support $\{0, 1, \ldots\}$ and parameter $r/(\lambda + r)$, denoting the number of Poisson arrivals with rate λ in an exponentially distributed interval of average length $1/r$. Let $h(s)$ be the Laplace transform of Y and $\theta = E[Y]$.

To fully parameterize equation (18) it remains to determine $h(s)$. To this end, let the random variable Y_f be defined as

$$Y_f = \max\{X_1, X_2, \ldots, X_f, Z\} \tag{59}$$

where $X_1, \ldots, X_f$ are exponential random variables with mean s/μ and Z is an exponential random variable with mean u. Conditioning on the number of queued customers and denoting by $h_f(s)$ the Laplace transform of Y_f,

$$\begin{aligned} h(s) &= \sum_{f=0}^{\infty} h(s|L=f)P(L=f) \quad (60) \\ &= \sum_{f=0}^{\infty} h_f(s)P(L=f) \quad (61) \end{aligned}$$

Our goal now is to compute $h_f(s)$. First, we present a closed formula expression for $E[Y_f]$,

$$E[Y_f] = \frac{u}{C(f+(1/u)(s/\mu), f)} + \frac{s}{\mu}\Big(\Gamma + \Psi(f+1)\Big) \tag{62}$$

where $\Gamma = 0.5772...$ is the Euler's constant, $\Psi(z)$ is the di-gamma function,

$$\Psi(z) = \frac{d}{dz}\ln\Gamma(z), \qquad \Gamma(z) = \int_{t=0}^{\infty} e^{-t}t^{z-1}dt \tag{63}$$

and $C(m,n)$ is the generalized binomial coefficient, $C(m,n) = \Gamma(m+1)/(\Gamma(n+1)\Gamma(m-n+1))$.

$E[Y_f]$ can be obtained noting that the cdf of Y_f is $F(x) = P(Y_f \le x) = (1 - e^{-\frac{\mu}{s}x})^f(1 - e^{-(1/u)x})$,

$$F(x) = (1 + \sum_{j=1}^{f}(-1)^j C(f,j)e^{-\frac{\mu}{s}jx})(1 - e^{-(1/u)x})$$

and $E[Y_f] = \int_{x=0}^{\infty}(1 - F(x))dx$, which yields (62) after some algebraic manipulation.

We now turn to the computation of the Laplace transform of Y_f. Let $W_{(i)}$ be the i^{th} order statistic among $\{X_1, \ldots, X_f, Z\}$. Consider the event

$$W_{(1)} \le \ldots \le W_{(i^*-1)} \le Z \le W_{(i^*+1)} \le \ldots \le W_{(f)}$$

We compute the Laplace transform of Y_f conditioning on the value of i^*,

$$h_f(s) = \sum_{j=1}^{f+1} P(i^* = j)h_f(s|i^* = j) \tag{64}$$

where $P(i^* = j)$ equals

$$\frac{1/u}{(f-j+1)(\mu/s) + 1/u}\prod_{i=1}^{j-1}\frac{(\mu/s)(f-i+1)}{(f-i+1)(\mu/s)+1/u}$$

Let U_i denote an exponential random variable with rate $(f-i+1)(\mu/s) + (1/u)$, $1 \le i \le i^*$, and V_i denote an exponential random variable with rate $i(\mu/s)$, $1 \le i \le f - i^* + 1$. Then

$$h_f(s|i^* = j) = E[e^{-s\max\{W_{(1)},\ldots,W_{(i^*)},\ldots,W_{(f)}\}}|i^* = j] =$$

$$= E[e^{-s(\sum_{i=1}^{j}U_i + \sum_{i=1}^{f-j+1}V_i)}] = \prod_{i=1}^{j}E[e^{-sU_i}]\prod_{i=1}^{f-j+1}E[e^{-sV_i}]$$

$$= \prod_{i=1}^{j}\frac{(f-i+1)(\mu/s)+(1/u)}{(f-i+1)(\mu/s)+(1/u)+s}\prod_{i=1}^{f-j+1}\frac{i(\mu/s)}{i(\mu/s)+s}$$

Substituting this last expression into (64) yields $h_f(s)$. □

Remark on the proof of lemma 3.2. Our goal now is to show how the average waiting time of a peer arriving in an idle period (equation (58)) differs from the average of the waiting times of all peers arriving in an idle period.

Consider one idle period. In that idle period, let L be the total number of requests queued up. Let $Q^{(j)}$ be the queueing time of the j^{th} request, $1 \le j \le L$. The average queueing time of a customer conditioned on the customer being queued is

$$E[Q|L>0] = \frac{E[\sum_{j=1}^{L}Q^{(j)}|L>0]}{E[L|L>0]} \tag{65}$$

where $E[L|L>0] = \frac{\lambda+r}{r}$ and

$$E[\sum_{j=1}^{L} Q^{(j)}|L>0] = \sum_{l=1}^{\infty} P(L=l|L>0)E[\sum_{j=1}^{L} Q^{(j)}|L=l]$$

$$= \sum_{l=1}^{\infty} r/(\lambda+r)(\lambda/(\lambda+r))^{l-1} \sum_{i=1}^{l} \frac{i}{\lambda+r}$$

$$= \sum_{l=1}^{\infty} r/(\lambda+r)(\lambda/(\lambda+r))^{l-1} \frac{l(l+1)}{2(\lambda+r)}$$

$$= \frac{\lambda+r}{r^2}$$

Therefore

$$E[Q|L>0] = \frac{1}{r} \tag{66}$$

Let the average of the waiting times of peers arriving in an idle period be denoted by Q',

$$E[Q'|L>0] = E\Big[\frac{\sum_{j=1}^{L} Q^{(j)}}{L}|L>0\Big] \tag{67}$$

Then,

$$E[Q'|L>0] = \sum_{l=1}^{\infty} P(L=l|L>0)E[Q'|L=l]$$

$$= \sum_{l=1}^{\infty} r/(\lambda+r)(\lambda/(\lambda+r))^{l-1} \frac{1}{l} \sum_{i=1}^{l} \frac{i}{\lambda+r}$$

$$= \sum_{l=1}^{\infty} r/(\lambda+r)(\lambda/(\lambda+r))^{l-1} \frac{l+1}{2(\lambda+r)}$$

$$= \frac{1}{2}\left(\frac{1}{r} + \frac{1}{\lambda+r}\right) \tag{68}$$

Equation (66) may be contrasted with (68). Even though the average wait time experienced by a peer arriving during an idle period is $1/r$ (equation (66)), the expected value of the average wait time in an idle period is smaller than or equal to $1/r$ (equation (68)).

End of remark.

Proof of theorem 3.2

In the following proof we assume $r_k = R$ for all $k = 1, \ldots, K$.

PROOF. Part (a) follows from the fact that $E[\mathcal{T}] = Ks/\mu + \mathcal{P}/R$, $E[T] = s/\mu + P/R$ and $\mathcal{P} < P$.

Part (b) follows since

- $\lim_{K\to\infty} \mathcal{P} = 0$ and

$$\lim_{K\to\infty} E[\mathcal{T}] = \frac{Ks}{\mu} + \frac{1}{R}\mathcal{P} = \frac{Ks}{\mu} \tag{69}$$

- $\lim_{R\to 0} P = 1$ and

$$\lim_{R\to 0} E[T] = \frac{s}{\mu} + \frac{1}{R}P = \frac{s}{\mu} + \frac{1}{R} \tag{70}$$

Assuming that the optimal bundle groups files that are of interest to all users downloading the bundle, [2]

$$\lim_{R\to 0}\lim_{K\to\infty} \frac{E[T]}{E[\mathcal{T}]/K} = \frac{\frac{s}{\mu} + \frac{1}{R}}{s/\mu} = \Theta\Big(\frac{1}{R}\Big) \tag{71}$$

Note: In the derivation above K and R are independent. Nevertheless, if $K = \sqrt{\log(1/R)}$ a similar result still holds,

$$\lim_{K\to\infty} \frac{E[T]}{E[\mathcal{T}]/K} = \frac{\frac{s}{\mu} + e^{K^2}}{\frac{s}{\mu} + \frac{e^{K^2}}{K} e^{-\Theta(K^2)}} = \tag{72}$$

$$= \Theta(K^b e^{cK^2}) = \Theta\Big(\frac{(\sqrt{\log(1/R)})^b}{R^c}\Big), b \geq 0, c > 0$$

where the first equality follows from theorem 3.1. □

Extension of theorem 3.2 If $K = \sqrt{\log(1/R)}$ then $E[T]/E[\mathcal{T}] = \Theta\Big(\frac{1}{R^c(\sqrt{\log(1/R)})^b}\Big)$, $b \geq 0, c > 0$.

PROOF. If $K = \sqrt{\log(1/R)}$,

$$\lim_{K\to\infty} \frac{E[T]}{E[\mathcal{T}]} = \frac{\frac{s}{\mu} + e^{K^2}}{K\frac{s}{\mu} + e^{K^2} e^{-\Theta(K^2)}} = \tag{73}$$

$$= \Theta(e^{cK^2}/K^b) = \Theta\Big(\frac{1}{R^c(\sqrt{\log(1/R)})^b}\Big), b \geq 0, c > 0$$

where the first equality follows from theorem 3.1. □

Proof of lemma 3.3

PROOF. Due to the memoryless property of the exponential random variable, the virtual customer that starts the residual busy period is characterized by a random variable $Y = \max\{X_1, \ldots, X_n\}$ where $X_1, \ldots, X_n$ are exponential random variables with mean s/μ. Therefore, Y is an hypoexponential distribution with parameters $(s/\mu, s/(2\mu), \ldots, s/(n\mu))$, which has Laplace transform $\prod_{i=1}^{n}(i\mu/s)/(s + i\mu/s)$ and mean $\sum_{i=1}^{n} \frac{s}{i\mu}$. Equation (18) can be used to compute $B(n,0)$ for any value of n. $B(n,0)$ equals

$$\Big[\sum_{i=1}^{n} \frac{s}{i\mu} + \sum_{i=1}^{\infty} \frac{(\lambda(s/\mu))^i (s/\mu)[1 - h(i/(s/\mu))]}{i!i}\Big]$$

which, after some algebraic manipulation, leads to (12). Let us denote by $T_{i,j}$ the time it takes for a residual busy

[2] Note that in the rest of this paper the download time of an individual content, $E[T]$, is compared against the download time of the bundle, $E[\mathcal{T}]$, rather than $E[\mathcal{T}]/K$. The former comparison is applicable if only one of the files inside the bundle is of interest to each user while the later is more adequate if all users downloading the bundle are interested in all files. In the remark following the proof of 3.2 we derive the download time theorem in case the metric of interest is $\frac{E[T]}{E[\mathcal{T}]}$ rather than $\frac{E[T]}{E[\mathcal{T}]/K}$.

period which starts with i peers to reach a population size of $j < i$ peers, where $B(i,j) = E[T_{i,j}]$. For $n > l$ and $n > k > l$, we have that $T_{n,l} = T_{n,k} + T_{k,l}$. Therefore, in general $E[T_{n,l}] = E[T_{n,k}] + E[T_{k,l}]$ and in particular, $E[T_{n,l}] = E[T_{n,0}] - E[T_{l,0}]$. Equation (12) and

$$B(n,l) = B(n,0) - B(l,0) \tag{74}$$

provide a way to compute $B(n,l)$ for arbitrary values of n and $l < n$. □

Proof of theorem 3.3

PROOF. The probability that a request leaves without being served, P, is

$$P = \frac{1/r}{E[B] + 1/r} \tag{75}$$

We now compute $E[B]$, the expected length of the busy period of a system which cycles between three phases: (1) one or more publishers are available; (2) no publisher is available but content is still available; (3) the content is not available.

We denote by $\overline{G}$ the expected length of a residual busy period starting when the number of publishers vanishes to zero. Assuming peers reached steady state before the number of publishers went to zero,[3]

$$\overline{G} = B(m) = \sum_{i=0}^{\infty} \frac{e^{-\frac{\lambda s}{\mu}} \left(\frac{\lambda s}{\mu}\right)^i}{i!} B(i,m) \tag{76}$$

The system starts in (1). It cycles between (1), (2), (1), (2) and so on up to reaching phase (3). Let us denote the number of times that it goes through (1) and (2), before reaching (1) for the last time and finally (2) and (3), by a geometrically distributed random variable C. C has support {0, 1, ... } and success probability p,

$$p = P\{X > \overline{G}\} = e^{-r\overline{G}} \tag{77}$$

where X is an exponentially distributed random variable with rate r. Therefore,

$$E[C] = \frac{1-p}{p} = \frac{1 - e^{-r\overline{G}}}{e^{-r\overline{G}}} \tag{78}$$

Phase (1) takes on average $\frac{e^{ru}-1}{r}$, the busy period of a population formed only by publishers. Phases (1) and (2) together, when not followed by (3), take on average

$$\frac{e^{ru}-1}{r} + E[X|X < \overline{G}] \tag{79}$$

where

$$E[X|X < \overline{G}] = \frac{\frac{1}{r} - \left(\frac{1}{r} + \overline{G}\right)e^{-r\overline{G}}}{1 - e^{-r\overline{G}}} \tag{80}$$

[3] $\overline{G}$ refers to the expected duration of the residual busy period whereas $E[B]$ refers to the expected duration of the busy period. In the main body of the article, $\overline{G}$ is refered to as $B(m)$.

The expected busy period is

$$E[B] = E[C]\left(\frac{e^{ru}-1}{r} + E[X|X < \overline{G}]\right) + \frac{e^{ru}-1}{r} + \overline{G} \tag{81}$$

Substituting (76), (78) and (80) into (81) yields

$$E[B] = \frac{e^{r(\overline{G}+u)} - 1}{r} \tag{82}$$

Replacing (82) into (75) leads to

$$P = \exp(-r(u + \overline{G})) \tag{83}$$

□

Proof of remark following theorem 3.3.

In the following proof we assume $r_k = r$ for all $k = 1, \ldots, K$.

PROOF. If $R = Kr$ and $U = Ku$ then, from (83),

$$\mathcal{P} = e^{-\Theta(K^2)} \tag{84}$$

therefore the availability theorem holds.

Assuming that the optimal bundle groups files that are of interest to all users downloading the bundle,

$$\lim_{r\to 0} \lim_{K\to\infty} \frac{E[T]}{E[\mathcal{T}]/K} = \frac{\frac{s}{\mu} + \frac{1}{r}\frac{1/r}{E[B]+1/r}}{\frac{s}{\mu} + \frac{1}{KR}\mathcal{P}} = \Theta\left(\frac{1}{r}\right) \tag{85}$$

In the derivation above r is constant when $K \to \infty$. Alternatively, if $\frac{1}{R} = e^{K^2}$ then $\frac{1}{r} = Ke^{K^2}$,

$$\lim_{K\to\infty} \frac{E[T]}{E[\mathcal{T}]/K} = \frac{\frac{s}{\mu} + \frac{K}{R}}{\frac{s}{\mu} + \frac{1}{KR}\mathcal{P}} = \frac{\frac{s}{\mu}e^{-K^2} + K}{\frac{s}{\mu}e^{-K^2} + \frac{1}{K}\mathcal{P}} =$$

$$= \Theta(K^b e^{cK^2}) = \Theta\left(\frac{(\sqrt{\log(1/R)})^b}{R^c}\right), b \geq 0, c > 0$$

which grows unbounded as $K \to \infty$ ($R \to 0$). □

Extension of remark following 3.3 If $K = \sqrt{\log(1/R)}$ then $E[T]/E[\mathcal{T}] = \Theta\left(\frac{(\sqrt{\log(1/R)})^b}{R^c}\right)$, $b \geq 0, c > 0$.

PROOF. If $K = \sqrt{\log(1/R)}$,

$$\lim_{K\to\infty} \frac{E[T]}{E[\mathcal{T}]} = \frac{\frac{s}{\mu} + Ke^{K^2}}{K\frac{s}{\mu} + e^{K^2}e^{-\Theta(K^2)}} = \tag{86}$$

$$= \Theta(K^b e^{cK^2}) = \Theta\left(\frac{(\sqrt{\log(1/R)})^b}{R^c}\right), b \geq 0, c > 0$$

where the first equality follows from theorem 3.1. □

Another remark on the proof of theorem 3.3.

If the Laplace transform $G^\star(s)$ of the residual busy period G is known then $E[B]$ is given by

$$E[C]\left(\frac{e^{ru}-1}{r} + E[X|X < G]\right) + \frac{e^{ru}-1}{r} + E[G|X > G] \tag{87}$$

Let us compute $E[X|X < G]$ and $E[G|X > G]$, in that order.

The probability of zero arrivals of a Poisson process with rate r in an interval G is

$$P\{X > G\} = \int_0^\infty e^{-rx} g(x)dx = G^\star(r) \tag{88}$$

Also,

$$E[X|X > G] = \int_{g=0}^\infty E[X|X > g]P\{g \le G \le g+dg|X > G\} \tag{89}$$

$$= \int_{g=0}^\infty \frac{(E[X] + g)P\{g \le G \le g + dg\}e^{-\lambda g}}{G^\star(r)} \tag{90}$$

$$= \frac{1}{r} - \frac{G'^\star(r)}{G^\star(r)} \tag{91}$$

Substituting (88) and (91) into $E[X]$,

$$\begin{aligned} E[X] &= E[X|X < G](1 - P\{X > G\}) + \\ &\quad +E[X|X > G]P\{X > G\} \end{aligned}$$

leads to

$$\begin{aligned} E[X] &= \frac{1}{r} = E[X|X < G](1 - G^\star(r)) + \qquad (92) \\ &\quad +\Big(\frac{1}{r} - \frac{G'^\star(r)}{G^\star(r)}\Big)G^\star(r) \end{aligned}$$

and

$$E[X|X < G] = \frac{\frac{1}{r} - (\frac{1}{r} - \frac{G'^\star(r)}{G^\star(r)})G^\star(r)}{1 - G^\star(r)} \tag{93}$$

$$E[X|X < G] = \frac{1}{r} + \frac{G'^\star(r)}{1 - G^\star(r)} \tag{94}$$

The last residual busy period in each busy period lasts for $E[G|X > G]$. Due to the memoryless property of the exponential distribution,

$$E[X] = E[X - G|X > G]$$

which together with (89) yields

$$E[G|X > G] = -\frac{G'^\star(r)}{G^\star(\lambda)} \tag{95}$$

Putting together (88)–(95) and (87) we get that $E[B]$ equals

$$\frac{1 - G^\star(r)}{G^\star(r)}\Big(\frac{e^{ru} - 1}{r} + \frac{1}{r} + \frac{G'^\star(r)}{1 - G^\star(r)}\Big) + \frac{e^{ru} - 1}{r} - \frac{G'^\star(r)}{G^\star(r)} \tag{96}$$

which after algebraic simplification yields

$$E[B] = \frac{(e^{ru}/G^\star(r)) - 1}{r} \tag{97}$$

End of remark.

Derivation of equation (16)

Proof. The derivation is similar to the one of theorem 3.3. With a slight abuse of notation, the expected busy period is, similarly to (81),

$$E[\mathcal{B}] = E[C]\Big(U + E[X|X < \overline{G}]\Big) + U + \overline{G} \tag{98}$$

where

$$\overline{G} = B(m) = \sum_{i=0}^\infty \frac{e^{-\frac{K^2\lambda_s}{\mu}}\left(\frac{K^2\lambda s}{\mu}\right)^i}{i!} B(i, m) \tag{99}$$

$B(i, m)$ is obtained from lemma 3.3 substituting λ by $K\lambda$ in (12), X is an exponentially distributed random variable with rate R and

$$E[X|X < \overline{G}] = \frac{\frac{1}{R} - \Big(\frac{1}{R} + \overline{G}\Big)e^{-R\overline{G}}}{1 - e^{-R\overline{G}}} \tag{100}$$

Substituting (100) and (78) into (98) yields

$$E[\mathcal{B}] = \frac{e^{r\overline{G}}(UR + 1) - 1}{R} \tag{101}$$

and equation (16) follows from (3) and (101). □

Altruistic lingering

The availability and download time theorems hold in the scenario with altruistic lingering since adding a constant to the residence time of each peer does not change the asymptotics of the system when $K \to \infty$.

To model altruistic lingering we use equation (17). Let S be an exponentially distributed random variable with mean u, representing the residence time of publishers. Let L be the sum of two exponentially distributed random variables with means $\beta_1 = s/\mu$ and $\beta_2 = 1/\gamma$ representing the residence time of peers after starting the download. The distribution of L is

$$\begin{aligned} P(L \le x) &= L(x) = \qquad (102) \\ &= \frac{(1/\beta_1)(e^{-\frac{1}{\beta_1}x} - 1) - (1/\beta_2)(e^{-\frac{1}{\beta_2}x} - 1)}{(1/\beta_2) - (1/\beta_1)} \end{aligned}$$

if $\beta_1 \neq \beta_2$ and Erlang(2, β_1) otherwise. A customer arriving in a busy period draws its residence time X as follows,

$$X = \begin{cases} L & \text{with probability } q_1' \\ S & \text{with probability } 1 - q_1' \end{cases} \tag{103}$$

where $q_1' = \lambda/(\lambda + r)$. Therefore, we parameterize (17) setting $\beta = \lambda + r$, $H(x) = 1 - e^{-ux}$, $\theta = u$, $G(x) = (1 - q_1')S(x) + q_1' L(x)$.

For the bundled swarm, $\beta = \Lambda + R$, $H(x) = 1 - e^{-Ux}$, $\theta = U$, $q_1' = \Lambda/(\Lambda + R)$, $S(x) = 1 - e{-}Ux$, $\beta_1 = S/\mu$ and $\beta_2 = 1/\gamma$. □

Trace driven arrivals

In this section we report results obtained using peer arrival patterns observed in real swarms. We selected two files from the 2008 Olympic game opening ceremony (which we refer to as files A and B). We scaled the arrival rates so that the swarms in isolation were not popular but as a bundle were popular enough to be self-sustaining. In this Planetlab experiment, we set publisher up and down times, U and D, to be exponentially distributed with mean of 500s and 1500s, respectively. Both files have size S=10 MB and we set the capacity of peers and publishers as 50KBps and 100KBps, respectively.

For each of the three scenarios that we considered [(1) bundle, (2) file A isolated and (3) file B isolated] we ran a Planetlab experiment for 12 hours. The download distributions for the three considered scenarios are shown in Figure 8. The dots mark the mean and the box plots show the 1st, 2nd and 3rd quartiles and, the 5 and 95 percentiles. Note that the mean and variance in the bundled case are significantly smaller than the ones for the two individual files, with an improvement in download time of 39% and 41%, respectively.

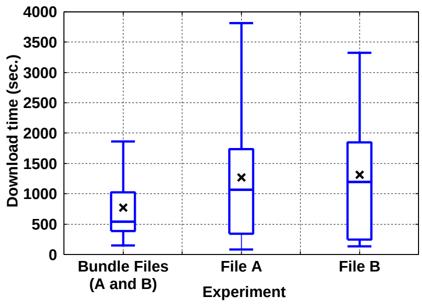

Figure 8: Experiment using trace-driven arrivals (Mean, quartiles, 5th and 95th percentiles)