On Sample-Path Staleness in Lazy Data Replication

Xiaoyong Li, *Student Member, IEEE*, Daren B. H. Cline, and Dmitri Loguinov, *Senior Member, IEEE*

***Abstract*—We analyze synchronization issues between two point processes, one modeling data churn at an information source and the other periodic downloads to its replica (e.g., search engine, web cache, distributed database). Due to pull-based synchronization, the replica experiences recurrent staleness, which translates into some form of penalty stemming from its reduced ability to perform consistent computation and/or provide up-to-date responses to customer requests. We model this system under non-Poisson update/refresh processes and obtain sample-path averages of various metrics of staleness cost, generalizing previous results and exposing novel problems in this field.**

***Index Terms*—Replication, staleness.**

I. Introduction

With the massive growth of the Internet and deployment of large-scale distributed applications, mankind faces new challenges in acquiring, processing, and maintaining vast amounts of data. In response to this flood of information, companies deploy cloud-based solutions designed to provide replicated and distributed support to the skyrocketing storage and processing demand of their users.

One interesting problem in these applications is the *highly-dynamic* nature of content, especially when perfect synchronization of sources, replicas, intermediate caches, and various computation is impossible. In fact, many large-scale distributed systems (e.g., airline reservations, online banking, web search engines, social networks) operate under constant data churn and may never see consistent snapshots of the entire network. As a result, these applications may hold and/or manipulate a mixture of objects that existed at the source at different times t in the past. This leads to questions about staleness, synchronization costs, and techniques for deciding optimal refresh policies.

In traditional databases, the source opens outbound communication with the replicas whenever it detects important changes. This enables *push-based* operation that actively expires stale content and broadcasts notifications into the system. Even under multi-hop replication, staleness lags in these systems are described by simple models that can be reduced to convolutions of single-hop notification delays. In other cases, however, scalability and administrative autonomy require that sources operate independently and provide information only based on explicit request, especially when they are unable to track their replicas or adopt modifications to existing protocols.

This *pull-based* replication (also called *optimistic* or *lazy*) improves both scalability of the service and availability of the data, but at the expense of increased age of manipulated content [25], [48]. This model of operation has enjoyed ubiquitous deployment in the current Internet (e.g., HTTP, DNS, network monitoring, web caching, RSS feeds, stock-ticker aggregators, certain types of CDNs, sensor networks); however, it still poses many fundamental modeling challenges. Our goal is to study them in this paper.

A. Motivation and Objectives

Suppose a *replica* is a system whose goal is to synchronize against *information sources*, apply certain processing to downloaded content, and serve results to *data consumers*. One challenge of this architecture is that sources not only require pull-based operation, but also lack the ability to predict future updates, which makes real-time estimates of remaining object lifetime (i.e., TTL) unavailable to the replica.

As information evolves at the source, which we call *data churn*, the replica may become stale and provide responses to that do not reflect the true state of the system. In such cases, we assume that user satisfaction and system performance are directly rated to the amount of time by which the replica is lagging behind the source. To convert time units into cost, suppose the application applies some weight function $w(x)$ to the age of stale content to determine the *penalty* associated with a particular refresh policy and data-churn process. Then, the goal of the system is to optimize the expectation of penalty observed by a stream of arriving customers.

This problem has been considered in the context of web systems [5], [6], [9]–[11], [13], [17], [18], [20], [23], [26], [29], [30], [33], [34], [38], [41], [47]; however, analytical results have predominantly assumed a Poisson update process at the source, with function $w(x)$ limited to either 1 or x. However, real systems driven by human behavior often require more complex families of processes (e.g., with heavy-tailed inter-update distributions, non-stationary dynamics, and slowly decaying correlation). Similarly, user sensitivity to outdated material may experience rapid increases for small x and eventual saturation for larger x, in which case other weight functions might be more appropriate. Since application performance under general update processes and wider classes of $w(x)$ is currently open, we aim to fill this void below.

Manuscript received February 18, 2015; revised August 19, 2015; accepted October 06, 2015; approved by IEEE/ACM Transactions on Networking Editor A. Wierman. Date of publication October 28, 2015; date of current version October 13, 2016. This work was supported by the NSF under Grants CNS-1017766 and CNS-1319984. An earlier version of the paper appeared in IEEE INFOCOM 2015.

X. Li and D. Loguinov are with the Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: xiaoyong@cse.com; dmitri@cse.tamu.edu).

D. B. H. Cline is with the Statistics Department, Texas A&M University, College Station, TX 77843 USA (e-mail: dcline@stat.tamu.edu).

Digital Object Identifier 10.1109/TNET.2015.2488595

B. Contributions

Consider a single source driven by an update process N_U and a single replica with the corresponding download process N_D, which is independent of N_U. Our first contribution is to propose a general framework for modeling staleness under arbitrary stochastic processes (N_U, N_D). Since staleness age and various penalties derived from it are usually defined in terms of sample-path averages [9]–[11], [13], [17], [18], [20], [29], [33], [34], [38], [41], [47], questions arise about their existence and possible variation across multiple realizations of the system. We address this issue by identifying the weakest set of conditions for which the distribution of staleness age exists and converges to a deterministic limit.

Armed with these results, our second contribution is to model interaction between the age processes of (N_U, N_D). For staleness to be a function of inter-update delays, we discover that sample-path ages of both processes, examined at the same random time Q_T, must be asymptotically *independent*. Interestingly, this condition does not automatically follow from independence of N_U and N_D, their stationarity, ergodicity, or even all three constraints combined. Instead, we show that it translates into a form of ASTA (Arrivals See Time Averages) [31], where the download process N_D must observe the sample-path distribution of update age.

Under the condition of age-independence, our third contribution is to derive the distribution of time by which the replica trails the source, the fraction of consumers that encounter a stale copy, the average number of missing updates from the replica at query time, and the general staleness cost under all suitable penalty functions $w(x)$. Our results involve simple closed-form expressions that are functions of limiting age distributions of both processes.

Our fourth contribution is to analyze conditions under which N_D produces provably optimal penalty for a given download rate. We show that penalty reduces if and only if inter-refresh delays become stochastically larger in second order. This leads to constant synchronization delays being optimal under all N_U and $w(x)$. This, however, presents problems in satisfying ASTA and creates a possibility of worst-case (i.e., 100%) staleness due to phase-lock between the source and the replica. To this end, we discuss broad requirements for ensuring that N_D avoids these drawbacks while remaining optimal.

We finish the paper with our last contribution that considers the practical aspects of staleness, including experimentation with Wikipedia page updates, error analysis of previous Poisson models, estimation of search-engine bandwidth requirements, and generalization to multiple sources/replicas.

II. Staleness Formulation

We start by explaining the underlying assumptions on the system, defining the various processes that determine information flow, and specifying the metrics of interest.

A. System Operation

We assume a model of distributed data generation, replication, and consumption shown in Fig. 1. During normal system operation, sources sustain random updates in response to external action (e.g., new posts on Facebook, traffic congestion in Google maps) or possibly some internal computation (e.g., MapReduce [15] indexing with periodic writes to disk). In either case, each update represents certain non-negligible information that manipulates the current state of the source.

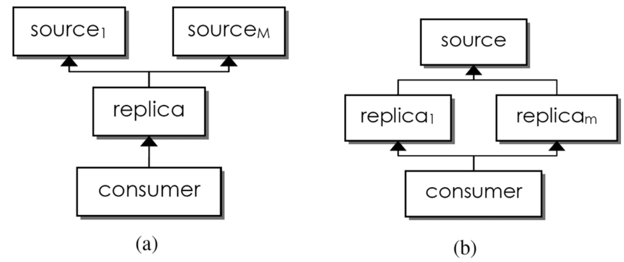


Fig. 1. System model (arrows signify the direction of information requests). (a) Aggregation. (b) Backup/load-balancing.

Replicas operate independently of the sources and perform one of the two general functions shown in the figure—many-to-one aggregation in part (a) and one-to-many replication in (b). The former case arises when the replica executes certain processing on multiple objects to provide the consumer with results that cannot be obtained otherwise. These applications include search engines, data-centric computing, and various web front-ends that cache queries against back-end databases. The purpose of the latter case is to handle failover during source crashes and/or ensure scalable load distribution under heavy customer demand. Applications in this category include CDNs, large websites, data centers (e.g., Amazon EC2), and various distributed file systems.

The final element of Fig. 1 is the consumer, which sends a stream of *requests* that represent either queries for information or attempts to recover the most-recent state of the source after it has crashed.

B. Updates and Synchronization

We next model interaction between a single source and a single replica, which is a prerequisite to understanding system performance. Suppose the source undergoes updates at random times $0 = u_1 < u_2 < \cdots$ and define $N_U(t) = \max\{i : u_i \le t\}$ to be a stochastic process that counts the number of updates in $[0, t]$. When referring to the entire process, rather than its value at some point, we omit t and write simply N_U.

For the replica, denote its random download (synchronization) instances by $0 = d_1 < d_2 < \cdots$ and the corresponding point process by $N_D(t) = \max\{k : d_k \le t\}$. This formulation neglects processing delays and treats all events as instantaneous. We additionally assume that both processes are simple and independent. Now, suppose the inter-update delays of N_U are given by a random process $\{U_i\}_{i=1}^{\infty}$ and inter-download cycles of N_D by $\{D_k\}_{k=1}^{\infty}$, which are illustrated in Fig. 2. Each of these sequences may be of fairly general nature, e.g., correlated and/or non-stationary.

C. Cost of Staleness

To understand the penalty of outdated content, suppose $M(t) = N_U(t) - N_U(d_{N_D(t)})$ counts the number of updates *missing* from the replica at time t (e.g., in Fig. 2, $M(t) = 2$).

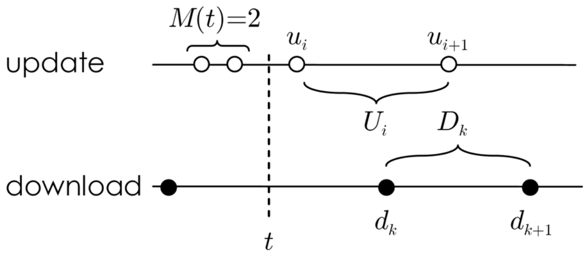


Fig. 2. Process notation.

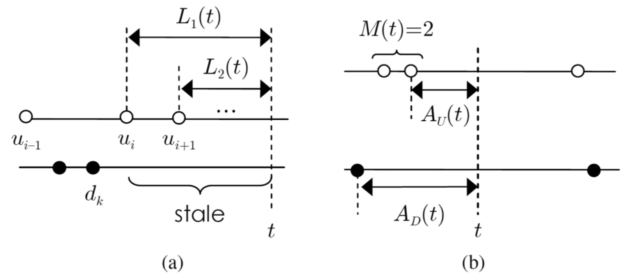


Fig. 3. Penalty lags and process age. (a) Staleness lags. (b) Age.

This is a discrete-state process that increments for each update and resets to zero for each synchronization.

Definition 1: A replica is called *stale* at time t if $M(t) > 0$. Otherwise, it is called *fresh*.

From the consumer's perspective, stale material reduces user satisfaction and lowers system performance, which needs to be translated into a cost metric that can be expressed via some known parameters of the system. The most basic penalty is the probability that the replica is stale at the time of request, i.e., $P(M(t) > 0)$, which determines how often users see outdated information and/or fail to fully restore a crashed source. The second obvious metric is the expected number of missing updates $E[M(t)]$, which measures the amount of lost information during a crash and estimates the difficulty in recreating it from the most recent checkpoint. This penalty is also important for Internet archiving applications that aim to capture every snapshot of the source [24] and situations when larger $M(t)$ may imply higher information divergence between the replica and the source.

More sophisticated cases are also possible. Suppose the source runs some computation, with updates representing certain intermediate states that are written to disk. A crash at time t requires computation to be restarted, which means that the penalty is determined not by $M(t)$, but rather by the *duration* of the computation that was lost due to staleness. Services that charge per CPU time-unit (e.g., Amazon EC2) may want to optimize against this metric rather than $E[M(t)]$. Furthermore, if the difficulty of recovering each update from other storage is proportional to the delay since the update was made, then staleness cost may be based on the *combined* lag of all missing updates at time t.

Definition 2: For a stale replica at time t, define lags $L_1(t) > L_2(t) > \cdots > L_{M(t)}(t)$ to be backward delays to each unseen update , i.e, $L_i(t) = t - u_{N_U(t)-M(t)+i}$.

This concept is illustrated in Fig. 3(a) for the first two lags. To keep the model general and cover the various options already seen in the literature [5], [6], [9], [10], we assume that the consumer is sensitive to either just lag $L_1(t)$, i.e., how long the *source* has been stale at time t, or the entire collection of lags $\{L_1(t), \ldots, L_{M(t)}(t)\}$, i.e., how long each uncaptured *update* has been stale. Since it is usually difficult to predict the value of information freshness to each customer, one requires a mapping from staleness lags to actual cost, which we assume is given by some non-negative weight function $w(x)$.

Definition 3: At time t, the *source penalty* is given by the weight of the delay since the replica was fresh last time:

$$\eta(t) = \begin{cases} w(L_1(t)) & M(t) > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{1}$$

while the *update penalty* is given by the aggregate weight of all staleness lags:

$$\rho(t) = \begin{cases} \sum_{i=1}^{M(t)} w(L_i(t)) & M(t) > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{2}$$

For example, $w(x) = 1$ produces the first two metrics discussed above, i.e., $P(M(t) > 0)$ via $E[\eta(t)]$ and $E[M(t)]$ via $E[\rho(t)]$. Both (1) and (2) are random variables, which suggests that system performance should be assessed by their average values. But as neither N_U nor N_D is assumed to be stationary, the expected penalty requires additional elaboration. Instead of considering $E[\eta(t)]$ and $E[\rho(t)]$, which may depend on time t, it is more natural to replace them with sample-path averages [10]:

$$\bar{\eta} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \eta(t)dt \text{ and } \bar{\rho} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \rho(t)dt, \tag{3}$$

where consumers are modeled as being equally likely to query the replica at any time in $[0, \infty)$.

D. Relationship to Prior Work

The majority of the literature on source penalty $\bar{\eta}$ is limited to Poisson N_U, either constant or exponential D, and $w(x) = 1$ or x [8]–[10], [13], [33], [42], [44], [47]. There has been only one attempt to model $\bar{\eta}$ under a renewal process N_U, in which [44] assumed $w(x) = 1$ and the entire sequence of refresh instances $\{d_1, d_2, \ldots\}$ was known. While appropriate in some cases, this model is difficult to evaluate in practice when N_D is given by its statistical properties.

Update penalty $\bar{\rho}$ has received less exposure, with almost all papers considering Poisson updates and just constant D. This includes $w(x) = 1$, where $\bar{\rho}$ is usually called *divergence* [23] or *blur* [16], with analysis available in [17], [18], [38], and $w(x) = x$, where $\bar{\rho}$ is known as *additive age* [28], *aggregated age* [29], *delay* [38], or simply *cost* [17]. Finally, $\bar{\rho}$ with a general $w(x)$ was called *obsolescence cost* in [18] and analyzed under a non-stationary Poisson N_U, but no closed-form results were obtained.

The Poisson assumption on N_U allows easy computation of the various metrics of interest. Outside these special cases, superposition of non-memoryless processes produces much more complex behavior.

III. Age Model

While (3) is convenient, it is unclear whether these limits exist, if they are finite, and under what conditions they are deterministic. We investigate these issues next.

A. Main Framework

We start by performing a convenient transformation of (3) to remove the integrals. Define Q_T to be a uniform random variable in $[0,T]$, which models the random query time of consumers. Suppose Q_T is independent of N_U and N_D, in which case (3) is the limit of $E[\eta(Q_T)|N_U,N_D]$ and $E[\rho(Q_T)|N_U,N_D]$ as $T \to \infty$. To keep formulas manageable, we sometimes omit explicit conditioning on processes (N_U,N_D); however, it should be noted that all expectations and probabilities are still computed within each sample path (i.e., with respect to Q_T only).

At time t, suppose age processes $A_U(t)$ and $A_D(t)$, shown in Fig. 3(b), specify delays to the previous update and synchronization event, respectively. Using this notation and observing that $M(t)>0$ is equivalent to $A_U(t)<A_D(t)$, define an ON/OFF staleness process:

$$S(t) = \begin{cases} 1 & A_U(t) < A_D(t) \\ 0 & \text{otherwise} \end{cases}, \tag{4}$$

whose properties at random time Q_T determine whether the consumer sees outdated information or not.

To analyze (4), our next topic is the behavior of $A_U(Q_T)$ and $A_D(Q_T)$ as $T \to \infty$, including existence of these limits and their relationship to $\{U_i\}_{i=1}^{\infty}$ and $\{D_k\}_{k=1}^{\infty}$.

Assumptions

We next aim to establish a minimal set of conditions under which analysis of staleness admits closed-form results. Consider a point process N with cycle lengths $\{X_i\}_{i=1}^{\infty}$, where each $X_i \sim F_i(x)$ is a random variable. In order for the age $A(Q_T)$ of this process to have a usable limiting distribution as $T \to \infty$, one must impose three constraints on N, which we discuss informally and motivate next, followed by a more rigorous definition.

The first restriction is that collection $\{X_i\}_{i=1}^{\infty}$ within each sample-path have some limiting distribution $F(x)$. If this fails to hold, staleness in (3) does not exist either. The second prerequisite is that $F(x)$ not be a random limit. This condition ensures that almost all sample-paths produce the same result. Finally, the third condition is that an $o(1)$ fraction of cycles must consume an $o(1)$ fraction of length as $n \to \infty$. Allowing otherwise would be a problem because $F(x)$, being a limiting distribution, does not capture these intervals, but Q_T still lands there with a non-diminishing probability as $T \to \infty$ (we discuss an example demonstrating this effect shortly).

Let $\mathbf{1}_A$ be an indicator variable of A and $\bar{F}(x) = 1 - F(x)$ the complementary CDF (cumulative distribution function) of $F(x)$. We are now ready to summarize our discussion.

Definition 4: A process N is called *age-measurable* if:

1) For all $x \geq 0$, except possibly points of discontinuity of the limit, sample-path distribution $H_n(x)$ of variables $\{X_1,\ldots,X_n\}$ converges in probability as $n \to \infty$:

$$H_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i \leq x} \xrightarrow{\text{P}} F(x). \tag{5}$$

2) Function $F(x)$ is deterministic with mean $0 < \delta < \infty$;

3) The average cycle length converges to δ in probability as $n \to \infty$:

$$Z_n := \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{P}} \delta = \int_0^{\infty} \bar{F}(x)dx. \tag{6}$$

Note that any renewal process $\{X_i\}_{i=1}^{\infty}$ satisfies this definition since all $F_i(x)$ are the same, which from the weak law of large numbers trivially leads to $F(x) = F_i(x)$ and $Z_n \to \delta$. Furthermore, condition (6) resembles mean-ergodicity, which is normally stated with a stronger type of convergence (e.g., mean-square or almost-sure) and only for *stationary* processes. For other cases, the fact that indicator variables are uniformly bounded allows application of the Dominated Convergence Theorem (DCT) [36] to show that $F(x)$ is the limiting average of individual distributions:

$$E[H_n(x)] = \frac{1}{n}\sum_{i=1}^{n} F_i(x) \to F(x). \tag{7}$$

It is pretty clear that (5) is not implied by (6). If $\{X_i\}_{i=1}^{\infty}$ are uniformly bounded, the reverse can be inferred (i.e., (6) follows from (5)); however, this does not hold universally. In fact, many random variables used in practice (e.g., exponential and Pareto) are not bounded and thus require an explicit assumption that convergence in (6) take place. Additionally, even if this limit exists, it does not generally equal δ, which is why we require that as well.

B. Distribution

Define the sample-path distribution of age $A(Q_T)$, in points Q_T uniformly placed in $[0,T]$, to be:

$$G(x,T) := P(A(Q_T) \leq x|N). \tag{8}$$

For an age-measurable process N, suppose the residual (age) distribution of its $F(x)$ is given by:

$$G(x) := \frac{1}{\delta}\int_0^x \bar{F}(y)dy. \tag{9}$$

It is well-known that a renewal [46] or regenerative [39] assumption on N yields $G(x,T) \to G(x)$ as $T \to \infty$. Our next result produces a condition that is both sufficient and necessary for this to hold.

Theorem 1: Process N is age-measurable if and only if $N(T)/T$ is almost surely bounded and $G(x,T)$ converges in probability to $G(x)$.

Proof: We start with the forward (sufficiency) proof. Consider an age-measurable N and let $S_k = \sum_{i=1}^{k} X_i$ be the k-th arrival point of this process. Note that almost-sure boundedness of $N(T)/T$ immediately follows from (6) and the fact that $\delta > 0$. The rest of the proof deals with convergence of $G(x,T)$.

Assume some constant $x \geq 0$. Then, event $A(Q_T) \leq x$ is equivalent to the existence of some $k \geq 1$ such that Q_T belongs to the k-th interval $[S_k, S_{k+1})$, under the condition that starting point $S_k \leq T$ and age $Q_T - S_k \leq x$. Defining

$$W_k = \min((T-S_k)^+, X_k, x), \tag{10}$$

where $(x)^+ = \max(x, 0)$, we get:

$$G(x,T) = \sum_{k=1}^{\infty} P(S_k \le Q_T < S_k + W_k | N). \quad (11)$$

Since Q_T is uniform in $[0, T]$, the probability that it falls into an interval of length W_k is simply W_k/T:

$$G(x,T) = \sum_{k=1}^{N(T)} \frac{\min(T - S_k, X_k, x)}{T}, \quad (12)$$

where the upper limit is reduced from ∞ to $N(T)$ since $(T - S_k)^+ = 0$ for $k > N(T)$. Recalling that all probabilities and expectations are dependent on the sample path, it follows that $G(x,T)$ is a random variable. Our goal below is to show it converges to a constant as $T \to \infty$. To this end, first observe that it can be bounded as:

$$\frac{\sum_{k=1}^{N(T)-1} \min(X_k, x)}{\sum_{k=1}^{N(T)} X_k} \le G(x,T) \le \frac{\sum_{k=1}^{N(T)} \min(X_k, x)}{\sum_{k=1}^{N(T)-1} X_k}, \quad (13)$$

where we use the fact that $T - S_k \ge X_k$ for all $k \le N(T) - 1$ and $T \in [\sum_{k=1}^{N(T)-1} X_k, \sum_{k=1}^{N(T)} X_k]$.

Next, notice that (5) implies that for all bounded, continuous functions $f(x)$ [36]:

$$\frac{1}{n} \sum_{i=1}^{n} f(X_i) \xrightarrow{\text{P}} \int_0^{\infty} f(x) dF(x), \quad (14)$$

which leads to:

$$\frac{1}{N(T)} \sum_{k=1}^{N(T)-1} \min(X_k, x) \xrightarrow{\text{P}} \int_0^{\infty} \min(y, x) dF(y). \quad (15)$$

Using (6), we also have:

$$\lim_{T \to \infty} \frac{1}{N(T)} \sum_{k=1}^{N(T)-1} X_k = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n-1} X_k = \delta. \quad (16)$$

Since both bounds in (13) have the same limit, $G(x,T)$ converges in probability[1] to the ratio of (15) to (16):

$$\frac{1}{\delta} \int_0^{\infty} \min(y, x) dF(y) = \frac{1}{\delta} \int_0^{x} \bar{F}(y) dy, \quad (17)$$

where the second integral follows from expanding the min function and integrating by parts.

We now present the reverse (necessity) proof. Assume that $G(x,T) \to G(x)$ for some $G(x)$ and $N(T)/T$ is bounded. Our goal is to show convergence of (5) and (6) to such deterministic limits that satisfy (9). From the Bolzano-Weierstrass theorem [19], every subsequence n_k contains a further subsequence $n'_k \to \infty$ such that

$$Z_{n'_k} \xrightarrow{\text{P}} \delta_{\{n'_k\}}, \quad (18)$$

where $\delta_{\{n'_k\}} > 0$ with probability 1 from $N(T)/T$ being bounded. Note that this limit may depend on the subsequence. To show finiteness of $\delta_{\{n'_k\}}$, notice that (13) and $G(x,T) \to G(x)$ implies that for all $x \ge 0$:

$$\frac{\sum_{i=1}^{n} \min(X_i, x)}{\sum_{i=1}^{n} X_i} \xrightarrow{\text{P}} G(x). \quad (19)$$

[1]Moreover, since $G(x,T)$ monotone and bounded by 1, convergence is uniform in x [3].

The left side of (19) is concave in $[0, \infty]$, which means that $G(x)$ must be either concave or degenerate at $x = 0$. From (18) and (19), we know that:

$$x \ge \frac{1}{n'_k} \sum_{i=1}^{n'_k} \min(X_i, x) \xrightarrow{\text{P}} \delta_{\{n'_k\}} G(x), \quad (20)$$

where letting $x \to 0$ establishes that the latter case is impossible. Therefore, $G(x)$ must be concave, $G(x) > 0$ for $x > 0$, and finally $\delta_{\{n'_k\}} < \infty$.

From Helly's selection theorem [3], there exists a further subsequence $n''_k \to \infty$ of n'_k along which (18) holds and:

$$H_{n''_k}(x) \xrightarrow{\text{P}} F_{\{n''_k\}}(x), \quad (21)$$

where the limit is a proper CDF from Prohorov's theorem [3]. What remains to prove is that limits (18), (21) are independent of the subsequence and establish their relationship to $G(x)$. Using an analog of (14) for subsequences and applying (15):

$$\frac{\sum_{i=1}^{n''_k} \min(X_i, x)}{\sum_{i=1}^{n''_k} X_i} \xrightarrow{\text{P}} \frac{1}{\delta_{\{n''_k\}}} \int_0^{x} \bar{F}_{\{n''_k\}}(y) dy. \quad (22)$$

Invoking (19), this limit equals $G(x)$. Assuming $F(x)$ is some CDF, a function can be represented in the form of (9) using a *unique* pair $(\delta, F(x))$. Therefore, $F(x)$ must be $F_{\{n''_k\}}(x)$, which shows that for every subsequence n_k there exists a further subsequence n''_k such that $H_{n''_k}(x) \to F(x)$ and $Z_{n''_k} \to \delta$. But this means that the full sequence $H_n(x) \to F(x)$ and $Z_n \to \delta \in (0, \infty)$, i.e., (5), (6), and (9) hold. ■

To elaborate on this result, consider independent variables:

$$X_i = \begin{cases} 1 & \text{wp } 1 - 1/\sqrt{i} \\ \sqrt{i} & \text{wp } 1/\sqrt{i} \end{cases}, \quad (23)$$

whose limiting distribution in (5) is a constant with $\delta = 1$. However, Z_n in (6) converges to 2. Consequently, the distribution of age $A(Q_T)$ cannot be determined based on $F(x)$. Even worse, (9) suggests the age is uniform in $[0, 1]$, while $A(Q_T)$ is asymptotically finite only with probability $1/2$.

C. Expectation

While Theorem 1 establishes when $A(Q_T)$ has a limiting distribution, convergence of expectation $E[A(Q_T)|N]$ or suitability of $G(x)$ for computing it are not guaranteed. Furthermore, given that consumers may apply generic weights $w(x)$ to the various age-related metrics, it is important to identify when $E[w(A(Q_T))|N]$ exists as $T \to \infty$.

To build intuition for the next result, assume $X \sim F(x)$ is a non-negative variable and define its age A to be a random variable with CDF $G(x)$ in (9). Then, we are interested in the relationship between $E[w(A)]$ and X. To this end, suppose for any locally integrable function $w(x)$, we set $w_1(x) = w(x)$ and then recursively integrate the result $n - 1$ times to define:

$$w_n(x) := \int_0^{x} w_{n-1}(y) dy. \quad (24)$$

Using integration by parts in Lebesgue-Stieltjes integrals and keeping in mind that $w_{n+1}(0) = 0$ for $n \ge 1$:

$$E[w_{n+1}(X)] = \int_0^{\infty} w_n(x) \bar{F}(x) dx = E[w_n(A)] E[X]. \quad (25)$$

Therefore, in order for $E[w(A)]$ to exist, one must ensure that both $E[w_2(X)]$ and $E[X]$ do. Note that the latter does so by (6), but the former requires an additional constraint.

Definition 5: A point process N is called *age-measurable by weight function* $w(x)$ if it is age-measurable and

$$\frac{1}{n}\sum_{i=1}^{n} w_2(X_i) \xrightarrow{P} \in_0^\infty w_2(x)dF(x) < \infty. \tag{26}$$

Note that age-measurable by a constant is equivalent to simply age-measurable since in that case (26) becomes (6). We omit the proof of the next result as it follows that of Theorem 1 pretty closely.

Theorem 2: For a process N that is age-measurable by $w(x)$, the sample-path expectation of $w(A(Q_T))$ converges in probability as $T \to \infty$:

$$\lim_{T\to\infty} E[w(A(Q_T))|N] = \int_0^\infty w(x)dG(x). \tag{27}$$

To understand this better, consider another counter-example:

$$X_i = \begin{cases} 1 & \text{wp } 1-1/i \\ 7\sqrt{i} & \text{wp } 1/i \end{cases}. \tag{28}$$

The limiting distribution in (5) is again a constant equal to 1, but this time (6) converges to $\delta = 1$, which makes this process age-measurable. However, for $w(x) = x$, the sum in (26) oscillates between 25 and 30 as n increases, while the corresponding integral is $1/2$. As a result, N is not measurable by $w(x)$ and $E[A(Q_T)]$ does not converge as $T \to \infty$.

From this point on, we omit explicit conditioning on the sample-path since results do not depend on N for age-measurable processes. However, we keep in mind that all probabilities and expectations involving Q_T are still taken in the sample-path sense.

IV. Staleness Cost

This section models the probability of staleness and expected cost under both penalty metrics defined earlier.

A. Age Independence

We now return to examining (4). In order to determine when the replica is stale, one requires comparison of $A_U(Q_T)$ with $A_D(Q_T)$, which may not be independent random variables, even if N_U and N_D are. To prevent such cases, which are called *phase-lock* [4], conditions known as ASTA (Arrivals See Time Averages) [31] must apply to the age of one process when sampled by the arrival points of the other. This issue is delayed until a later section, but now we define more clearly what independence of $A_U(Q_T)$ and $A_D(Q_T)$ means.

Specifically, suppose (N_U, N_D) are age-measurable. Then, let $F_U(x)$ and $F_D(x)$ be respectively the limiting CDF functions of interval lengths defined in (5), with the corresponding average rates μ and λ, i.e.,

$$\frac{1}{\mu} = \int_0^\infty \bar{F}_U(x)dx \quad \text{and} \quad \frac{1}{\lambda} = \int_0^\infty \bar{F}_D(x)dx. \tag{29}$$

Further, let $U \sim F_U(x)$ and $D \sim F_D(x)$ be random update and download cycle lengths. Similarly, suppose $G_U(x)$ and $G_D(x)$ are the limiting CDFs of age from (9), with lower-case functions $g_U(x)$ and $g_D(x)$ representing the corresponding PDFs. When the necessary limits exist, let $A_U \sim G_U(x)$ and $A_D \sim G_D(x)$ denote the two random ages as $T \to \infty$.

Definition 6: Two age-measurable point processes N_U and N_D are called *age-independent* if for all $x, y \geq 0$:

$$\lim_{T\to\infty} P(A_D(Q_T) < x | A_U(Q_T) = y) = G_D(x). \tag{30}$$

If either N_U or N_D is Poisson, (30) is guaranteed from PASTA (Poisson Arrivals See Time Averages) [45], which explains why prior work did not encounter these nuances. To shed more light on this condition, consider independent stationary renewal processes N_U and N_D with lattice[2] inter-arrival distributions, each with integer span. Due to stationarity, the first cycles are $U_1 \sim G_U(x)$ and $D_1 \sim G_D(x)$. Both $A_U(Q_T)$ and $A_D(Q_T)$ have continuous distributions; however, conditioning on the pair of sample paths, $A_U(Q_T) - A_D(Q_T) - U_1 + D_1$ can only take integer values. Consequently, age-independence (30) cannot hold, not even asymptotically.

Unconditionally, however, the ages are independent:

$$\begin{aligned} P(A_U(Q_T) \leq x, A_D(Q_T) \leq y) &= E[E[\mathbf{1}_{A_U(t)\leq x}\mathbf{1}_{A_D(t)\leq y}|Q_T]] \\ &= \frac{1}{T}\int_0^T E[\mathbf{1}_{A_U(t)\leq x}\mathbf{1}_{A_D(t)\leq y}]dt \\ &= \frac{1}{T}\int_0^T P(A_U(t)\leq x)P(A_D(t)\leq y)dt \\ &= G_U(x)G_D(y), \end{aligned} \tag{31}$$

although this has no bearing on staleness.

To give a more concrete meaning to the conditional probability in (30), we have the next result.

Theorem 3: Age-independence implies that $A_D(t)$ sampled in update points of N_U produces a sequence of random variables that converges in distribution to $G_D(x)$:

$$\lim_{T\to\infty} \frac{1}{N_U(T)} \sum_{i=1}^{N_U(T)} \mathbf{1}_{A_D(u_i)<x} = G_D(x). \tag{32}$$

Proof: First define $d(y,T)$ to be the number of points in $[0,T]$ where $A_U(t) = y$ occurs. Recalling that $u_1 = 0$, this can be expressed as:

$$d(y,T) = \sum_{i=1}^{N_U(T)} \mathbf{1}_{U_i\geq y, u_i+y\leq T} = \sum_{i=1}^{N_U(T)} \mathbf{1}_{U_i\geq y}\cdot\mathbf{1}_{u_i+y\leq T}.$$

Next, let $c(y,T)$ be the number of these points in which the download age A_D is smaller than x:

$$\begin{aligned} c(y,T) &= \sum_{i=1}^{N_U(T)} \mathbf{1}_{U_i\geq y, u_i+y\leq T, A_D(u_i+y)<x} \\ &= \sum_{i=1}^{N_U(T)} \mathbf{1}_{U_i\geq y}\cdot\mathbf{1}_{u_i+y\leq T}\cdot\mathbf{1}_{A_D(u_i+y)<x}. \end{aligned} \tag{33}$$

Noticing that

$$P(A_D(Q_T) \leq x | A_U(Q_T) = y) = \frac{c(y,T)}{d(y,T)} \tag{34}$$

[2]A random variable X is called *lattice* if there exists a constant c such that X/c is always an integer.

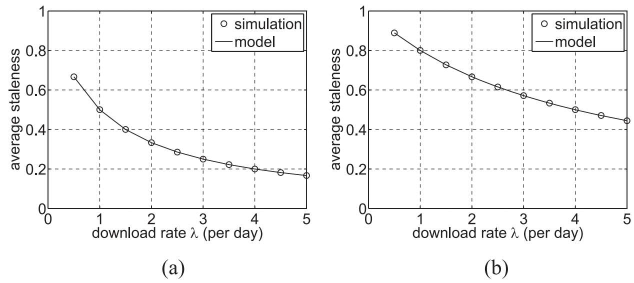


Fig. 4. Examination of (36) under $\mu = 2$. (a) Pareto U, constant D. (b) Constant U, Pareto D.

and applying Theorem 1, we get using (30):

$$G_D(x) = \lim_{T\to\infty} \lim_{y\to 0} \frac{c(y,T)}{d(y,T)} = \lim_{T\to\infty} \frac{1}{N_U(T)} \sum_{i=1}^{N_U(T)} \mathbf{1}_{A_D(u_i)<x}, \tag{35}$$

which is (32) with the two sides swapped. ∎

B. Preliminaries

Our first objective is to derive the probability of staleness.

Theorem 4: Assuming that N_U and N_D are age-independent, the probability of staleness at time Q_T converges in probability as $T \to \infty$ to:

$$P(S(Q_T) = 1) \to p := \mu \int_0^\infty \bar{F}_U(y)\bar{G}_D(y)dy. \tag{36}$$

Proof: Due to the existence and independence of $A_U(Q_T)$ and $A_D(Q_T)$ in the limit, we immediately obtain:

$$p = P(A_D > A_U) = \int_0^\infty \bar{G}_D(y)dG_U(y). \tag{37}$$

Expanding $dG_U(y) = \mu\bar{F}_U(y)dy$ leads to the result. ∎

To perform a self-check against prior results with Poisson N_U, observe that (36) simplifies to $p = 1 - \lambda(1 - e^{-\mu/\lambda})/\mu$ under constant D and $\mu/(\mu + \lambda)$ under exponential D, which are consistent with [10], [13]. Simulations in Fig. 4 examine model accuracy in more interesting cases of general renewal processes. We use Pareto CDF $1 - (1 + x/\beta)^{-\alpha}$ with $\alpha = 3$ and mean $\beta/(\alpha - 1) = \beta/2$. Observe in the figure that the model matches simulations very well, with constant download intervals performing significantly better against Pareto update cycles in (a) than the other way around in (b). For example, synchronizing pages at their update rate (i.e., $\lambda = \mu = 2$) serves stale copies with probability 33% in the former case and 66% in the latter. Furthermore, for the same p, the scenario in (a) requires roughly 4 times less bandwidth than in (b).

The next intermediate result is the distribution of the first lag $L_1(Q_T)$, which relies on p in (36).

Theorem 5: If N_U and N_D are age-independent, the CDF of $L_1(Q_T)$ converges in probability as $T \to \infty$ to:

$$F_L(x) = 1 - \frac{\mu}{p}\int_0^\infty \bar{F}_U(y)\bar{G}_D(x + y)dy. \tag{38}$$

Proof: Consider the ON/OFF staleness process in Fig. 5(a) and suppose the query time t falls in the ON period. Then, since t is uniformly random within this cycle, the backward delay $L_1(t)$ is symmetrical to the forward (residual) delay $R_D(t)$, meaning they have the same distribution. Note that it is important to condition on $A_D(t) > A_U(t)$ since residual $R_D(t)$ depends on age $A_D(t)$, i.e.,

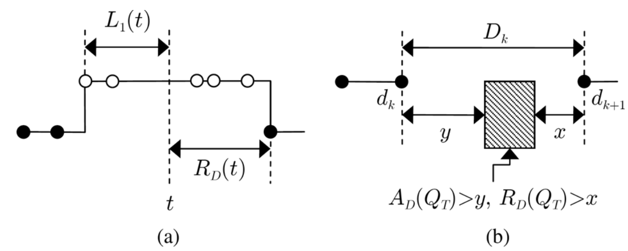


Fig. 5. Visualizing the proof of Theorem 5. (a) Staleness process $S(t)$. (b) Reward of each cycle.

$$P(L_1(t) > x) = P(R_D(t) > x|A_D(t) > A_U(t)). \tag{39}$$

Since N_U and N_D are age-independent, we can condition on $A_U(Q_T) = y$ without impacting the distribution of $A_D(Q_T)$ or $R_D(Q_T)$. Following the proof of Theorem 1, define $L_k = d_k + y$ and $M_k = \min(T, d_{k+1} - x)$ to be the lower/upper boundaries within synchronization interval k such that if $Q_T \in [L_k, M_k]$, then $R_D(Q_T) > x$ and $A_D(Q_T) > y$. See Fig. 5(b) for an illustration.

Define $C_T = P(L_1(Q_T) > x|A_U(Q_T) = y)$ and observe that it converges as $T \to \infty$:

$$\begin{aligned} C_T &= \frac{1}{pT}\sum_{k=1}^{\infty} P(L_k \le Q_T \le M_k) \\ &= \frac{1}{pT}\sum_{k=1}^{N(T)} (D_k - (x + y))^+ \\ &\to \frac{\lambda}{p}\int_0^\infty \max(z - (x + y), 0)dF_D(z) \\ &= -\frac{\lambda}{p}\int_{x+y}^\infty (z - (x + y))d\bar{F}_D(z) = \frac{\lambda}{p}\int_{x+y}^\infty \bar{F}_D(z)dz \\ &= \frac{1}{p}\int_{x+y}^\infty g_D(z)dz = \frac{\bar{G}_D(x + y)}{p}. \end{aligned} \tag{40}$$

Unconditioning $A_U(Q_T)$ and keeping in mind that its distribution is well-defined as $T \to \infty$, we get (38). ∎

Theorem 5 allows a simple expression for the fraction of requests $c(\tau)$ that observe content outdated by less than τ time units, which was called β-currency in [5] and Δ-consistency in [42]. This can be expressed as:

$$c(\tau) = 1 - \bar{F}_L(\tau)p = \int_0^\infty g_U(y)G_D(\tau + y))dy, \tag{41}$$

which conveniently simplifies to $P(A_D - A_U < \tau)$, where $A_D - A_U$ is the generalized lag between the replica and the source, i.e., non-positive values mean fresh states. Fig. 6 compares (41) to simulations using $\lambda = \mu$. As seen in the figure, this page retrieved at a random time is stale by less than $\tau = 0.4$ days (9.6 hours) with probability $c(\tau) = 98\%$ in the first case and 62% in the second.

C. Source Penalty

We are now ready to derive a general formula for $\bar{\eta}$.

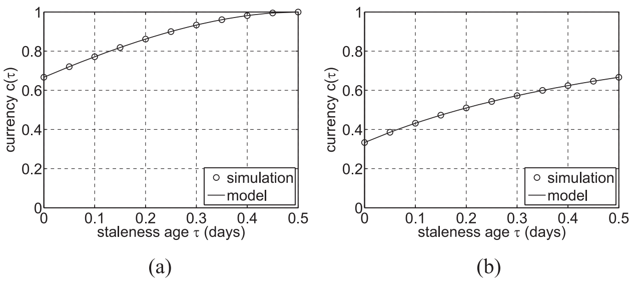


Fig. 6. Examination of (41) under $\lambda = \mu = 2$. (a) Pareto U, constant D. (b) Constant U, Pareto D.

Theorem 6: If N_U and N_D are age-independent, while N_D is age-measurable by $w(x)$, the source penalty converges in probability to:

$$\bar{\eta} = \lambda\mu \int_0^\infty \bar{F}_U(y) \int_0^\infty w(x)\bar{F}_D(x+y)dxdy. \quad (42)$$

Proof: First, observe that

$$\bar{\eta} = E[w(L_1)]P(A_D > A_U), \quad (43)$$

where $L_1 \sim F_L(x)$. Working back from (38), the tail CDF of L_1 can be written more compactly as:

$$P(L_1 > x) = P(A_D - A_U > x | A_D > A_U). \quad (44)$$

Since $L_1 = A_D - A_U > 0$, conditioned on $A_D > A_U$, it suffices that only N_D be measurable by $w(x)$. In that case:

$$\bar{\eta} = E[w(L_1)]P(A_D > A_U) = p\int_0^\infty w(x)dF_L(x), \quad (45)$$

or equivalently:

$$\bar{\eta} = \int_0^\infty g_U(y) \int_0^\infty w(x)g_D(x+y)dxdy \quad (46)$$

which immediately leads to (42) after expansion of $g_U(x) = \mu\bar{F}_U(x)$ and $g_D(x+y) = \lambda\bar{F}_D(x+y)$. ■

With $w(x) = 1$, (42) reduces to staleness probability p already discussed above. For the other case $w(x) = x$ seen in the literature, we obtain the expected *staleness age* $\bar{\eta} = E[L_1(t)]p$ by which the replica trails the source. Under Poisson N_U and constant D, we get from (42):

$$\bar{\eta} = \frac{1}{2\lambda} - \frac{1}{\mu} + \frac{\lambda(1-e^{-\mu/\lambda})}{\mu^2}, \quad (47)$$

and when both distributions are exponential:

$$\bar{\eta} = \frac{\mu}{\lambda(\lambda+\mu)}. \quad (48)$$

These special cases are consistent with [9]. Simulations in Fig. 7 additionally confirm that (42) is accurate under general renewal processes. Also observe in the figure that the combination in (b) continues to offer inferior performance to that in (a); however, the difference between the two scenarios is now more pronounced. For example, using the same $\lambda = \mu$ considered earlier, search-engine clients encounter indexing results outdated on average by 0.06 days (1.5 hours) in the left subfigure and by 0.8 days (19 hours) in the right. This example shows how drastically the cost changes based on the *shape* of $F_U(x)$ and $F_D(x)$,

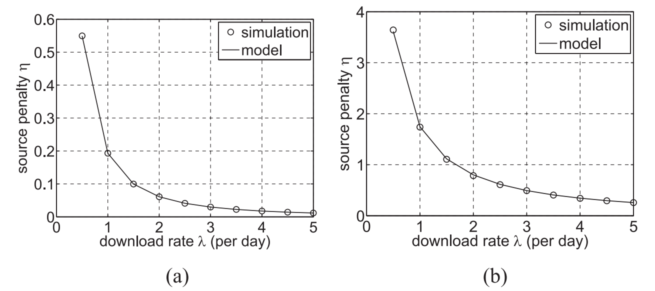


Fig. 7. Examination of (42) under $w(x) = x, \mu = 2$. (a) Pareto U, constant D. (b) Constant U, Pareto D.

which emphasizes the importance of utilizing models that can accurately handle any underlying processes (N_U, N_D).

We now offer a more intuitive look at source penalty. Modifying $w(x)$ to be zero for negative x, we can rewrite (42) in a more compact form:

$$\bar{\eta} = E[w(A_D - A_U)] = \lambda E[w_2(D - A_U)]. \quad (49)$$

This result shows that $\bar{\eta}$ is determined by the *positive* deviation of the generalized lag $A_D - A_U$ from zero, or equivalently by that of $D - A_U$, where the weight applied to each deviation is given respectively by $w(x)$ and $w_2(x)$. The only caveat is that simplification (49) requires weight functions that can explicitly handle negative arguments, e.g., a constant penalty would be $w(x) = \mathbf{1}_{x\geq 0}$ rather than just $w(x) = 1$. Throughout the rest of the paper, we avoid the extra notation dealing with $x < 0$, but keep this in mind.

D. Update Penalty

Unlike the previous section, we next show that $\bar{\rho}$ admits a much simpler result that depends only on the mean update rate μ rather than the entire distribution $F_U(x)$. This was first observed through simulations in [17] for constant D, but no explanation or extension to other cases was offered.

Theorem 7: Assuming N_U and N_D are age-independent, while N_D is age-measurable by $w_2(x)$, the update penalty converges in probability to:

$$\bar{\rho} = \mu E[w_2(A_D)] = \lambda\mu E[w_3(D)]. \quad (50)$$

Proof: Using Lebesgue-Stieltjes integrals and treating point processes as random measures, we can re-write (2) as:

$$\rho(t) = \int_{t-A_D(t)}^t w(t-s)dN_U(s). \quad (51)$$

Taking the expectation along each sample path:

$$\begin{aligned}\bar{\rho} &= \lim_{T\to\infty} E\left[\int_{Q_T - A_D(Q_T)}^{Q_T} w(Q_T - s)dN_U(s)\right] \\ &= \lim_{T\to\infty} \frac{1}{T}\int_0^T \int_{t-A_D(t)}^t w(t-s)dN_U(s)dt \\ &= \lim_{T\to\infty} \frac{1}{T}\int_0^T w_2(A_D(t))dN_U(t) \\ &= \lim_{T\to\infty} \frac{1}{T}\sum_{i=1}^{N_U(T)} w_2(A_D(u_i)). \quad (52)\end{aligned}$$

Applying (32), the sequence $\{A_D(u_1), A_D(u_2), \ldots\}$ sampled in update points $\{u_i\}$ converges in distribution to that of $A_D(Q_T)$ as $T \to \infty$. Then, (52) becomes:

$$\bar{\rho} = \lim_{T\to\infty} \mu E[w_2(A_D(Q_T))]. \tag{53}$$

Since N_D is $w_2(x)$-measurable, (27) shows that this expectation converges and its limit equals $\mu E[w_2(A_D)]$. By (25), this is also $\lambda\mu E[w_3(D)]$. ■

To compare against prior results, consider Poisson N_U and constant D. Then, (50) produces $\bar{\rho} = \mu/(2\lambda)$ for $w(x) = 1$ and $\mu/(6\lambda^2)$ for $w(x) = x$, both of which match previous analysis of these special cases [28], [38], [47]. Generalizing to exponential D, we obtain from (50) respectively μ/λ and μ/λ^2. Interestingly, this shows that switching downloads from constant intervals to exponential *doubles* the number of missing updates and *sextuples* their combined age.

For $w(x) = 1$, a simple closed-form expression is possible for all D:

$$E[M(t)] = \frac{\lambda\mu E[D^2]}{2} = \frac{\mu}{2\lambda}\left(1 + \lambda^2 Var[D]\right). \tag{54}$$

For example, Pareto D produces in (54):

$$E[M(t)] = \frac{\mu(\alpha-1)}{\lambda(\alpha-2)}, \tag{55}$$

which for $\alpha = 3$ is quadruple that of constant D and double that of exponential D. Another peculiar case is $\alpha \to 2$, where $E[M(t)]$ tends to infinity regardless of N_U. In fact, the update process itself may exhibit $Var[U] = \infty$, but the expected number of updates by which the replica falls behind will still become unbounded as α approaches 2.

Since source penalty $\bar{\rho}$ sums up the ages of all missing updates, it allows usage of *decaying* functions $w(x)$ such that their integral is increasing. We demonstrate this effect using $w(x) = 1/(1 + x)$, for which $w_2(x) = \log(1 + x)$. This cost function increases rapidly for small x, but then becomes less sensitive to staleness as the age of replicated content grows. Since $w_3(x) = (1 + x)\log(1 + x) - x$, constant D yields:

$$\bar{\rho} = \mu[(\lambda + 1)\log(1 + 1/\lambda) - 1]. \tag{56}$$

For $D \sim$ Pareto(α, β) with $\alpha = 3$ and $\beta = 2/\lambda$, we get:

$$\bar{\rho} = 2\mu \begin{cases} \frac{2\log(2/\lambda)-2+\lambda}{(\lambda-2)^2} & \lambda \neq 2 \\ 0.25 & \lambda = 2 \end{cases}. \tag{57}$$

Fig. 8 confirms that both models are accurate, with constant D enjoying a 60% lower penalty compared to Pareto.

V. Optimality

Motivated by (54) and consistently worse performance of Pareto D, the goal of this section is to understand the impact, if any, of $Var[D]$ on penalty and determine whether there exists an optimal distribution $F_D(x)$ that, for a fixed download budget λ, provably results in the lowest cost for all N_U and all suitable functions $w(x)$.

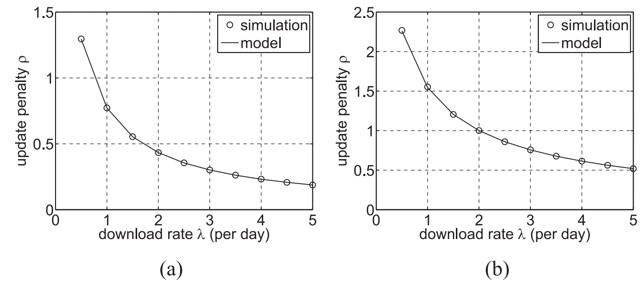


Fig. 8. Examination of (56) and (57) under $\mu = 2$. (a) Pareto U, constant D (b) constant U, Pareto D.

A. Stochastic Dominance

We start with general concepts from economics and game theory that are useful for understanding optimality. For two non-negative random variables $X \sim F_X(x)$ and $Y \sim F_Y(x)$, let their CDF difference be:

$$H(x) = F_Y(x) - F_X(x), \tag{58}$$

whose generalization $H_n(x)$ is given by (24). Then, we have the following definition.

Definition 7: Variable X is said to stochastically dominate Y in n-th order, which we write as $X \geq_{st}^n Y$, if $H_n(x) \geq 0$ for all $x \in \mathbb{R}$.

This concept is important because desirable characteristics of D can be inferred from those of A_D, as shown next.

Lemma 1: Assume $E[X] = E[Y]$ and $n \geq 2$. Then, X stochastically dominates Y in n-th order, i.e., $X \geq_{st}^n Y$, iff the age of Y stochastically dominates the age of X in $(n-1)$-st order, i.e., $A_Y \geq_{st}^{n-1} A_X$.

Proof: Let $G_X(x)$ and $G_Y(x)$ be the CDF of A_X and A_Y, respectively. Define:

$$J(x) = G_Y(x) - G_X(x), \tag{59}$$

which can be expressed using $H_2(x)$ as:

$$\begin{aligned} J(x) &= \frac{\int_0^x (1 - F_Y(x))dy}{E[Y]} - \frac{\int_0^x (1 - F_X(y))dy}{E[X]} \\ &= \frac{\int_0^x (F_X(y) - F_Y(y))dy}{E[X]} = -\frac{1}{E[X]}H_2(x). \end{aligned} \tag{60}$$

Integrating both sides $n - 2$ additional times leads to:

$$J_{n-1}(x) = -\frac{1}{E[X]}H_n(x). \tag{61}$$

From this and Definition 7, it follows that $X \geq_{st}^n Y$ implies $A_Y \geq_{st}^{n-1} A_X$ and vice versa. ■

As given by the next lemma, first-order stochastic dominance allows one to determine the relationship between expected utilities $E[w(X)]$ and $E[w(Y)]$. While it is possible to establish a more general version of this result using n-th order dominance, it would restrict $w(x)$ to a narrower class of functions and thus would be less useful in practice.

Lemma 2: Condition $X \geq_{st}^1 Y$ holds iff for all non-decreasing functions $w(x)$ it follows that $E[w(X)] \geq E[w(Y)]$.

B. Penalty Analysis

Returning to the topic of information staleness, our goal is to determine the condition under which both types of penalty can be reduced without changing the refresh rate. Define $\bar{\eta}(D_1)$ and $\bar{\eta}(D_2)$ to be the source penalties corresponding to random synchronization intervals D_1 and D_2, both with mean $1/\lambda$. For the opposite problem, i.e., finding the worst update distribution, define $\bar{\eta}(U_1)$ and $\bar{\eta}(U_2)$ to be the penalties that correspond to update intervals U_1 and U_2 under a fixed μ.

The next result shows that stochastic (rather than variance) ordering is needed to improve staleness penalty. Define $w(x)$ to be a *measure* if it is non-negative, non-decreasing, and right-continuous with $w(x)=0$ for $x<0$.

Theorem 8: Assume the conditions of Theorem 6. For a given N_U and fixed download rate λ, $D_1 \geq_{st}^2 D_2$ iff $\bar{\eta}(D_1) \leq \bar{\eta}(D_2)$ for all measures $w(x)$. Similarly, with a given N_D and fixed μ, $U_1 \geq_{st}^2 U_2$ iff $\bar{\eta}(U_1) \geq \bar{\eta}(U_2)$ for all measures $w(x)$.

Proof: Using (49), observe that $\bar{\eta} = E[w(A_D - A_U)]$ is fully determined by the properties of variable $X = A_D - A_U$. For a fixed A_U, it is not difficult to show that X becomes stochastically smaller in first order iff A_D does. Applying Lemmas 1–2, this means that penalty $\bar{\eta}$ gets smaller iff D increases stochastically in second order.

Similarly, for a fixed A_D, X gets stochastically larger in first order iff A_U becomes stochastically smaller. Again applying Lemmas 1–2, penalty $\bar{\eta}$ increases iff U becomes stochastically larger in second order. ■

A similar result holds under update penalty $\bar{\rho}$. Note that all $F_U(x)$ with the same μ are equivalent here, which is why we state only half of Theorem 8, and $w(x)$ is less restricted.

Theorem 9: Assume the conditions of Theorem 7. For a given N_U and fixed λ, $D_1 \geq_{st}^2 D_2$ iff $\bar{\rho}(D_1) \leq \bar{\rho}(D_2)$ for all non-negative $w(x)$.

Proof: Since $\bar{\rho} = \mu E[w_2(A_D)]$, where $w_2(x)$ is a measure for all non-negative $w(x)$, Lemmas 1–2 yield that $\bar{\rho}$ decreases iff D gets stochastically larger in second order. ■

The preceding results set up motivation to ask the question of whether there exists a distribution that dominates all others in second order. We answer this next.

Lemma 3: For a given mean, a constant stochastically dominates all other random variables in second order.

Proof: Suppose l is the fixed mean of all distributions under consideration. Let $F_X(x) = \mathbf{1}_{x>l}$ be the CDF of a constant and $F_Y(x)$ be the CDF of another random variable Y such that $E[Y]=l$. Our goal is to show that $H_2(x) \geq 0$.

When $x \leq l$, we have trivially:

$$H_2(x) = \int_0^x (F_Y(y) - F_X(y))dy = \int_0^x F_Y(y)dy \geq 0. \quad (62)$$

For $x > l$, we get:

$$\begin{aligned} H_2(x) &= \int_0^l F_Y(y)dy + \int_l^x (F_Y(y)-1)dy \\ &= l + \int_0^x F_Y(y)dy - x = l - \int_0^x (1-F_Y(y))dy \\ &\geq l - \int_0^\infty (1-F_Y(y))dy = 0, \end{aligned} \quad (63)$$

where we use the fact that $l = \int_0^\infty (1-F_Y(y))dy$. ■

This leads to the main result of this section.

Theorem 10: When the conditions of Theorems 8–9 hold, constant inter-synchronization delays are optimal under the corresponding staleness metric.

This allow us to resolve the relationship between the variance of D and penalty. If $E[D_1]=E[D_2]$, then $D_1 \geq_{st}^2 D_2$ implies $Var[D_1] \leq Var[D_2]$, but the opposite is not true. This shows that for a given download rate, just reducing the variance of refresh intervals, without enforcing $D_1 \geq_{st}^2 D_2$, is *insufficient* to improve the penalty across *all* functions $w(x)$. As an example, recall the special case of $\bar{\rho}$ with $w(x)=1$ in (54), where the penalty was reduced iff the variance of D was; however, no such causality exists for $w(x)=x$ or $\log(1+x)$. On the other hand, if reduction in penalty holds for all measures $w(x)$, then stochastic ordering between D_1 and D_2 follows and thus variance has to decrease (i.e., ordering of variances is *necessary*, but not sufficient).

C. Phase-Lock

Even though constant D is optimal from the staleness perspective, it unfortunately fails to guarantee age-independence (30) against *all* underlying N_U. We now deal with principles related to ASTA (Arrivals See Time Averages) [31], placing them in our context. In general, ASTA can be viewed as a condition that allows discrete and continuous sample-path averages of a process $X(t)$ to be equal almost surely:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X(t_k) = \lim_{T\to\infty} \frac{1}{T}\int_0^T X(t)dt. \quad (64)$$

Let $X(t) = \mathbf{1}_{A_D(t)<x}$ and $t_k = u_k$. Then, if (64) holds for all x, it follows that the distribution of refresh age $A_D(t)$ sampled in update points u_k equals that sampled in uniformly random instances Q_T, which in turn is equivalent to our earlier formulation (30). While we have given conditions for the right side of (64) to exist and equal a constant almost surely, existence of the left side or its equality to the integral is not guaranteed. ASTA analysis focuses on the properties of points $\{t_k\}$ and their relationship to $X(t)$ that allow (64) to hold; however, this normally requires conditions that are difficult to verify in practice (e.g., LAA, WLAA, LBA [31]). We therefore discuss guidelines for ensuring that (64) is satisfied, without becoming engrossed in unnecessary rigor.

While sampling constant update cycles with constant synchronization intervals sometimes leads to phase-lock, we next discuss how to achieve asymptotic age-independence in such cases. To build intuition, suppose $U_i = 5$ and $D_k = \pi$ for all $k \geq 1$. Then, from the equidistribution theorem, $A_U(d_k) = k\pi (\text{mod } 5)$ is a uniformly random variable in $[0,5]$, meaning that $A_U(d_k)$ has the same distribution as $A_U(Q_T)$. The key observation is to ensure that N_D puts its download points uniformly across the cycles of N_U.

In general, sequence $a_k = k\xi (\text{mod } T)$, where T is rational, ξ is irrational, and $k \in \mathbb{N}$, is uniformly distributed in $[0,T]$, in which case for any Riemann-integrable function, an ASTA-like condition automatically holds:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n f(a_k) = \frac{1}{T}\int_0^T f(x)dx. \quad (65)$$

While using $D_k = \pi$ to sample $U_i = 5$ works well, there is a possibility that U_i itself happens to be a multiple of π. To preclude these cases, N_D must exhibit enough randomness to prevent D_k/U_i from becoming deterministically an integer. One option for doing so is to require that either process employ non-lattice cycle lengths. Recall that non-lattice distributions may be entirely continuous, including the classical PASTA (Poisson Arrivals See Time Averages) [45] and the uniform distribution often suggested for network measurement [4]. However, they can also be entirely discrete. In such cases, cycle lengths must distribute mass across at least two values (a, b), where a/b is irrational, e.g., pairs $(\pi, 1)$ or $(e, \sqrt{2})$. By bringing spread $|b - a|$ closer to zero, it is possible to obtain a variety of approximations to the optimal (constant) synchronization delay with mean $l = (a + b)/2$.

Note that a non-lattice distribution may still enter phase-lock if its cycle lengths follow a deterministic pattern, e.g., both updates and downloads strictly alternate between 1 and π. To rule out these cases, it is sufficient to require that the non-lattice process randomize its delays, leaving the other one general. This produces the following.

Theorem 11: If either N_U or N_D uses iid, non-lattice cycle lengths, age-independence holds.

Proof: The result follows from the equidistribution theorem and the iid nature of delays. ∎

VI. Applications

We now examine the presence of Poisson updates in real data sources and show how to apply the developed models to solve several classes of multi-source/replica problems.

A. Real-Life Update Processes

We first discuss possible reasons for the frequent use of memoryless source-update processes in the literature. If indeed this is universal, extensions to non-Poisson dynamics may be unnecessary. While modeling convenience is one possible explanation [13], there is certain belief in the field that updates to individual web pages can be accurately described by a Poisson process, which has fueled this line of modeling for over a decade [5], [6], [9]–[11], [17], [18], [20], [23], [26], [29], [30], [33], [34], [38], [41], [47].

Intuitively, there is no fundamental reason why a single source should exhibit Poisson dynamics, especially when modified by humans. A more likely scenario would be heavy-tailed behavior observed in many areas of computer networks [14], [27], [35] and user-driven distributed systems [7], [37], [40]. Another intuitively reasonable inter-update distribution is constant, where certain information is injected into the system periodically by design or is obtained from an ON/OFF source (e.g., sensors trying to conserve energy).

Closer examination of the origin [10] of the Poisson conclusion reveals several limitations. First, the distribution of page inter-update intervals was sampled using *incomplete observation*, meaning that some of the updates went unnoticed. As a result, bias could have been introduced in the measurements. Second, the exponential distribution was fitted to updates of *multiple* pages rather than a single page. Poisson dynamics have been known to emerge when aggregating arrival processes [1] and summing up variables [2], which does not tell us much about the individual distributions being combined. Finally, to conclude that N_U is Poisson, it is insufficient to observe an exponential distribution in $\{U_i\}_{i=1}^{\infty}$; instead, one must also show stationary independent increments [46].

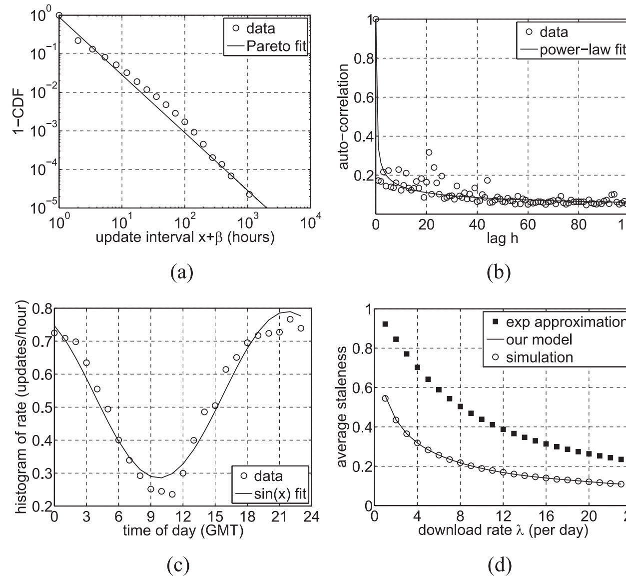


Fig. 9. George W. Bush page dynamics. (a) Distribution tail $\bar{F}_U(x)$. (b) Correlation $\rho(h)$. (c) Update rate. (d) Model accuracy.

B. Wikipedia

Even though certain measurement studies [5], [12], [21], [32] have found non-Poisson updates among web pages, they also lack ground truth. These pitfalls can be avoided if model verification is performed over sources that expose information about *each* update. One particularly interesting source with public traces of all modification timestamps is Wikipedia [43]. From a search-engine perspective, this website represents a realistic example of data churn stemming from user interaction with each other (e.g., edits from other people), flash crowds in response to external events, and diurnal activity patterns of the human lifecycle. Wikipedia is also well-suited for purposes of model validation and discussion.

To shed light on the complexity of real $F_U(x)$, we plot in Fig. 9(a) the tail CDF of inter-update delay for the most frequently modified article – "George W. Bush" with 44,296 updates in 10 years (mean delay $E[U] = 1.86$ hours). The figure is a close match to Pareto tail $(1 + x/\beta)^{-\alpha}$ with $\alpha = 1.4$ and $\beta = 0.93$. In Fig. 9(b), we show the corresponding auto-correlation function $\rho(h)$ with a power-law fit $h^{-0.37}$, which suggests long-range dependence (LRD) with Hurst parameter 0.81. Of course, LRD effects might be caused and/or compounded by non-stationarity. To address this question, Fig. 9(c) shows the update rate throughout the day, clearly indicating non-stationary dynamics.

This example underscores the need to keep the model general and not limit results to renewal or even stationary cases, which was our goal with assumptions (5)–(6). Approximating $F_U(x)$ as non-lattice and using constant D, we next compute the probability of staleness for this page by supplying (42) with George

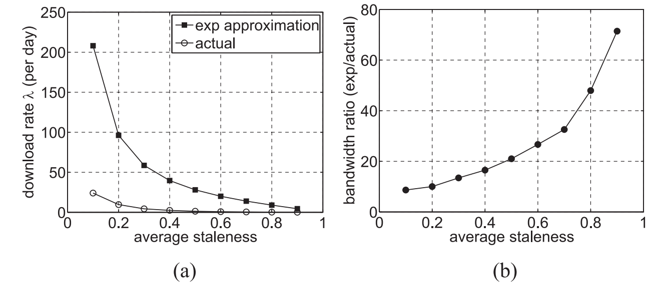


Fig. 10. Application of staleness models to the update process of George W. Bush. (a) Crawl rate. (b) Ratio.

W. Bush' empirically computed distribution $F_U(x)$. We contrast the result against the closest Poisson formula $1-\lambda(1-e^{-\mu/\lambda})/\mu$ from [10]. Fig. 9(d) shows that (42) is accurate, but the Poisson approximation suffers over 100% relative error for much of the examined range.

What is more important is the performance of the model in providing an accurate assessment of the download bandwidth needed to achieve a given p. We invert the formulas to solve for λ as a function of p and plot the result in Fig. 10(a). These results show a much more dramatic difference. For example, 20% staleness requires 95 downloads/day according to previous Poisson models, while in reality this can be achieved with just 8. To illustrate this better, we show the ratio of these two curves in Fig. 10(b), where the amount of Poisson overestimation varies from one to almost two orders of magnitude depending on the desired p.

C. Aggregation (Many-to-One)

When a single replica tracks M sources, as in Fig. 1(a), performance is assessed by its ability to provide usable aggregate information to the consumer. If sources are independent, many results are relatively easy to obtain. For example, consider a system that selects a replica and loads it with a MapReduce job that has to execute over the data of all sources. A computation may be considered successful if at least one source is fresh at the time of job request. Then, the fraction of successful attempts is $1-\prod_{i=1}^{M} p_i$, where p_i in (36) is the probability of staleness for source i. Alternatively, application consistency may require that *all* sources be simultaneously fresh, which leads to the probability of success via $\prod_{i=1}^{M}(1-p_i)$.

A more interesting problem is optimal allocation of download rates to different sources. Suppose q_i is the probability that an incoming query requests data from source i and μ_i is its update rate. Then, the goal is to allocate refresh rates λ_i so as to optimize the expected staleness cost $C(\lambda_i, \mu_i)$ for a given bandwidth budget Λ:

$$\min \sum_{i=1}^{M} q_i C(\lambda_i, \mu_i) \qquad \text{subject to} \qquad \sum_{i=1}^{M} \lambda_i \le \Lambda, \quad (66)$$

where $C(\lambda_i, \mu_i)$ refers to either $\bar{\eta}$ or $\bar{\rho}$.

For $\Lambda \ll \sum_{i=1}^{M} \mu_i$ and certain choices of $w(x)$, solutions to (66) using cost $\bar{\eta}$ are known to completely starve frequently modified sources in favor of those that are updating slowly [10]. Since (66) does not have a closed-form solution under $\bar{\eta}$ even in the simplest cases, specific conditions for starvation are not clear. Complete loss of synchronization for sources whose μ_i is above some (typically unknown) threshold may be an unwelcome surprise for many applications. This naturally leads to the question of whether $\bar{\rho}$ suffers from the same drawback. We address this next.

Theorem 12: Assume $q_i\mu_i > q_j\mu_j > 0$ and let refresh delays be optimal (i.e., constant). Then, the solution to (66) using $\bar{\rho}$ guarantees that $\lambda_i > \lambda_j > 0$.

Proof: Using Lagrange multipliers, we get that all partial derivatives of $q_i C(\mu_i, \lambda_i)$ must equal some constant κ:

$$\kappa = -\frac{\partial[q_i C(\mu_i, \lambda_i)]}{\partial \lambda_i} = -q_i\mu_i \frac{\partial[\lambda_i w_3(1/\lambda_i)]}{\partial \lambda_i}, \quad (67)$$

which follows from (50) and $D_i = 1/\lambda_i$. Expanding, we get $\kappa = q_i\mu_i f(\lambda_i)$, where

$$f(x) = \frac{w_2(1/x)}{x} - w_3(1/x) \quad (68)$$

is a monotonically non-increasing function:

$$f'(x) = -\frac{w_2(1/x)}{x^2} - \frac{w(1/x)}{x^3} + \frac{w_2(1/x)}{x^2} = -\frac{w(1/x)}{x^3}. \quad (69)$$

Notice that $f(\lambda_i) \ge 0$ for all λ_i since w_3 is an integral of $w_2(x)$ from 0 to $1/\lambda_i$. Therefore, $\kappa \ge 0$ and the relationship between μ_i and λ_i is determined by:

$$\lambda_i = f^{-1}\left(\frac{\kappa}{q_i\mu_i}\right). \quad (70)$$

If since f is non-increasing, larger $q_i\mu_i$ implies larger λ_i. Finally, since $f(x) > 0$ for all $x > 0$, it follows that its inverse f^{-1} has the same properly and thus no positive $q_i\mu_i > 0$ can achieve $\lambda_i = 0$. This means the optimal allocated rate must be strictly positive (i.e., no starvation). ■

To explain how optimization with $\bar{\rho}$ can be used, we assume constant D and $w(x) = 1$, with the goal to maximize $\sum_{i=1}^{M} q_i E[M_i(t)]$. Solving (66), the optimal download rate of each page is proportional to the square root of $q_i\mu_i$:

$$\lambda_i = \Lambda \frac{\sqrt{q_i\mu_i}}{\sum_{j=1}^{M}\sqrt{q_j\mu_j}}. \quad (71)$$

The optimal penalty is then:

$$\sum_{i=1}^{M} q_i E[M_i(t)] = \sum_{i=1}^{M} \frac{q_i\mu_i}{2\lambda_i} = \frac{\left(\sum_{j=1}^{M}\sqrt{q_j\mu_j}\right)^2}{2\Lambda}. \quad (72)$$

Define random variable μ to have the same distribution as $\{\mu_1, \dots, \mu_M\}$. Then, for the most basic scenario where all pages are equally popular, i.e., $q_i = 1/M$, we get:

$$\sum_{i=1}^{M} q_i E[M_i(t)] = M\frac{E[\sqrt{\mu}]^2}{2\Lambda}. \quad (73)$$

For the other extreme, where pages are searched for in proportion to their modification rate, i.e., $q_i \sim \mu_i$, we have:

$$\sum_{i=1}^{M} q_i E[M_i(t)] = M\frac{E[\mu]}{2\Lambda}. \quad (74)$$

To put these models in perspective, we use Wikipedia's distribution of μ, which happens to be quite heavy-tailed (i.e., Zipf shape $\alpha = 0.6$). The average update rate across all pages is $E[\mu] = 8$ updates/day; however, 98% of them exhibit μ_i less than 1/day, 90% less than 1/week, and 50% below 8/year. Using this distribution in (73) and (74) shows that optimizing staleness of the entire Wikipedia under uniform page access $q_i = 1/M$ requires 46 times less bandwidth Λ than under Zipf. This can be explained by the fact that keeping frequently modified pages fresh costs more bandwidth. This effect is related to the variance of $\sqrt{\mu}$:

$$\frac{E[\mu]}{E[\sqrt{\mu}]^2} = \frac{Var[\sqrt{\mu}] + E[\sqrt{\mu}]^2}{E[\sqrt{\mu}]^2} = \frac{Var[\sqrt{\mu}]}{E[\sqrt{\mu}]^2} + 1. \quad (75)$$

Consider extrapolating these results to $M = 100$ B sources and keeping the expected consumer lag $\sum_{i=1}^{M} q_i E[M_i(t)]$ below ω updates. We use the two models above as lower/upper bounds on the actual search-engine crawl rate. The first case requires download capability $\Lambda_1 = M \cdot E[\sqrt{\mu}]^2/2\omega = 99/\omega$ thousand pages per second (pps), while the second one $\Lambda_2 = M \cdot E[\mu]/2\omega = 4.6/\omega$ million pps. For $\omega = 10$ and 25 KB per page, these translate into 2 and 92 Gbps, respectively. Results can be easily adjusted to non-Wikipedia situations as long as $E[\sqrt{\mu}]$ and $E[\mu]$ are known.

D. Load-Balancing (One-to-Many)

The issue of redundant replication from a single source, as in Fig. 1(b), to m nodes is quite different from the opposite case considered in the previous subsection. When the source fails, suppose the goal is to deduce the expected penalty afforded by the freshest member of the entire collection of m replicas. The issue at stake is how this $1 \times m$ case compares to a single replica with some refresh rate λ and optimal D. To keep comparison fair, assume that each of the m replicas is allowed budget λ/m in synchronization with the source. Decentralized operation leads to much better robustness under failure, but is it possible that this causes reduced freshness? If so, what is the amount of extra download bandwidth needed to keep both scenarios equally stale?

The main caveat in solving this problem is that staleness at different replicas is no longer independent. This happens because updates at the source simultaneously make all copies outdated, which means that reliability does not benefit exponentially with increased m. To overcome this issue, let $N_D^1, \dots, N_D^m$ be the download processes used by the individual replicas. Then, observe that the entire collection can be replaced by a single replica that implements a refresh pattern N_D^*, which is a *superposition* of all point processes $\{N_D^i\}_{i=1}^m$. Therefore, the source can be recovered during the crash with a probability determined solely by N_D^*.

If we assume centralized scheduling between the replicas, then it is possible to run the system optimally (i.e., using a perfectly spaced out round-robin) and thus keep the overall penalty exactly the same as with a single replica. Under fully decentralized (i.e., independent) replica operation and $m \to \infty$, each rate $\lambda/m \to 0$ and thus N_D^* likely converges in distribution to a Poisson process with rate λ (Palm-Khintchine theorem [22]). This creates a problem, however, because exponential D requires noticeably more overhead than constant D to achieve the same staleness penalty. For example, using our model for $\bar{\rho}$ and discussion after (54), this difference is by a factor of 2 for $w(x) = 1$ and by a factor of 6 for $w(x) = x$, which shows that a distributed cluster of replicas may need to consume 100–500% more bandwidth than a centralized solution for a given level of QoS (quality-of-service).

E. Many-to-Many

We conclude the paper by noting that Internet applications often combine the last two scenarios, i.e., $M \times 1$ and $1 \times m$ replication, into a single framework. However, these problems are usually separable into subproblems that can be reduced to the analysis above. For example, suppose we are interested in the probability that a query to a random subset of j replicas finds at least one of the k sources fresh. First, we compute the staleness probability for each source based on the aggregate synchronization process N_D^* from j replicas. Second, since each source is independent, we multiply these probabilities to deduce the likelihood that all k sources are stale. Taking the complement of the result, we get the desired probability.

VII. Conclusion

The paper introduced a novel model of sampled age under general non-Poisson update/synchronization processes and applied it to obtain many useful metrics of staleness. We additionally established that constant inter-refresh intervals were optimal for all considered cases and provided guidelines for achieving ASTA even in those cases. We finally considered a family of related problems stemming from $1 \times m$ and $M \times 1$ replication, showing that they can be easily solved from the preceding analysis of the 1×1 case.

Future work involves reducing staleness when N_D is allowed to depend on observations of N_U and/or prior knowledge of its distribution of update cycles $F_U(x)$.

References

[1] S. L. Albin, "On Poisson approximations for superposition arrival processes in queues," *Manage. Sci.*, vol. 28, no. 2, pp. 126–137, 1982.

[2] R. Arratia, L. Goldstein, and L. Gordon, "Two moments suffice for Poisson approximations: The Chen-Stein method," *Ann. Probab.*, vol. 17, no. 1, pp. 9–25, Jan. 1989.

[3] R. B. Ash and C. A. Doleans-Dade, *Probability & Measure Theory*, 2nd ed. New York, NY, USA: Academic Press, 1999.

[4] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot, "The role of PASTA in network measurement," in *Proc. ACM SIGCOMM*, Sep. 2006, pp. 231–242.

[5] B. E. Brewington and G. Cybenko, "How dynamic is the web," *Comput. Netw.*, no. 1–6, pp. 257–276, Jun. 2000.

[6] L. Bright, A. Gal, and L. Raschid, "Adaptive pull-based policies for wide area data delivery," *Trans. Database Syst.*, no. 2, pp. 631–671, Jun. 2006.

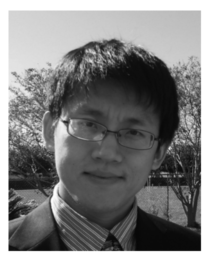
[7] F. E. Bustamante and Y. Qiao, "Friendships that last: Peer lifespan and its role in P2P protocols," in *Web Content Caching and Distribution*. Norwell, MA, USA: Kluwer, Sep. 2003.

[8] D. Carney, S. Lee, and S. Zdonik, "Scalable application-aware data freshening," in *Proc. IEEE ICDE*, Mar. 2003, pp. 481–492.

[9] J. Cho and H. Garcia-Molina, "The evolution of the web and implications for an incremental crawler," in *Proc. VLDB*, Sep. 2000, pp. 200–209.

[10] J. Cho and H. Garcia-molina, "Synchronizing a database to improve freshness," in *Proc. ACM SIGMOD*, May 2000, pp. 117–128.

[11] J. Cho and H. Garcia-Molina, "Estimating frequency of change," *Trans. Internet Technol.*, vol. 3, pp. 256–290, Aug. 2003.
[12] G. L. Ciampaglia and A. Vancheri, "Empirical analysis of user participation in online communities: The case of Wikipedia," in *Proc. ICWSM*, Sep. 2010, pp. 219–222.
[13] E. G. Coffman, Z. Liu, and R. R. Weber, "Optimal robot scheduling for web search engines," *J. Scheduling*, no. 1, pp. 15–29, Jun. 1998.
[14] M. E. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: Evidence and possible causes," *IEEE/ACM Trans. Netw.*, vol. 5, no. 6, pp. 835–846, Dec. 1997.
[15] J. Dean and S. Ghemawat, "MapReduce: Simplified data processing on large clusters," in *Proc. USENIX OSDI*, Dec. 2004, pp. 137–150.
[16] D. Denev, A. Mazeika, M. Spaniol, and G. Weikum, "SHARC: Framework for quality-conscious web archiving," in *Proc. VLDB*, Aug. 2009, vol. 2, no. 1, pp. 586–597.
[17] D. Dey, Z. Zhang, and P. De, "Optimal synchronization policies for data warehouses," *INFORMS J. Comput.*, no. 2, pp. 229–242, Jan. 2006.
[18] J. Eckstein, A. Gal, and S. Reiner, "Monitoring an information source under a politeness constraint," *INFORMS J. Comput.*, no. 1, pp. 3–20, Jan. 2008.
[19] P. M. Fitzpatrick, *Advanced Calculus*, 2nd ed. Belmont, CA, USA: Thomson Brooks/Cole, 2006.
[20] A. Gal and J. Eckstein, "Managing periodically updated data in relational databases: A stochastic modeling approach," *J. ACM*, vol. 48, no. 6, pp. 1141–1183, Nov. 2001.
[21] D. Gruhl *et al.*, "How to build a webfountain: An architecture for very large-scale text analytics," *IBM Syst. J.*, vol. 43, no. 1, pp. 64–77, 2004.
[22] D. Heyman and M. Sobel, *Stochastic Models in Operations Research*. New York, NY, USA: McGraw-Hill, 1982, vol. 1.
[23] Y. Huang, R. H. Sloan, and O. Wolfson, "Divergence caching in client-server architectures," in *Proc. IEEE PDIS*, Sep. 1994, pp. 131–139.
[24] "Internet Archive," [Online]. Available: http://archive.org/
[25] R. Ladin, B. Liskov, L. Shrira, and S. Ghemawat, "Providing high availability using lazy replication," *Trans. Comput. Syst.*, no. 4, pp. 360–391, Nov. 1992.
[26] J.-J. Lee, K.-Y. Whang, B. S. Lee, and J.-W. Chang, "An update-risk based approach to TTL estimation in web caching," in *Proc. WISE*, Dec. 2002, pp. 21–29.
[27] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic," in *Proc. ACM SIGCOMM*, Sep. 1993, pp. 183–193.
[28] Y. Ling and W. Chen, "Measuring cache freshness by additive age," *Oper. Syst. Rev.*, vol. 38, pp. 12–17, Jul. 2004.
[29] Y. Ling and J. Mi, "An optimal trade-OFF between content freshness and refresh cost," *Appl. Probab.*, vol. 41, no. 3, pp. 721–734, Sep. 2004.
[30] N. Matloff, "Estimation of Internet file-access/modification rates from indirect data," *Trans. Model. Comput. Simul.*, vol. 15, pp. 233–253, Jul. 2005.
[31] B. Melamed and W. Whitt, "On arrivals that see time averages," *Oper. Res.*, vol. 38, no. 1, pp. 156–172, 1990.
[32] A. Ntoulas, J. Cho, and C. Olston, "What's new on the web? the evolution of the web from a search engine perspective," in *Proc. WWW*, May 2004, pp. 1–12.
[33] C. Olston and J. Widom, "Best-effort cache synchronization with source cooperation," in *Proc. ACM SIGMOD*, May 2002, pp. 73–84.
[34] C. Olston and S. Pandey, "Recrawl scheduling based on information longevity," in *Proc. WWW*, Apr. 2008, pp. 437–446.
[35] K. Park, G. Kim, and M. Crovella, "On the relationship between file sizes, transport protocols, and self-similar network traffic," in *Proc. IEEE ICNP*, Oct. 1996, pp. 171–180.
[36] S. Resnick, *A Probability Path*. Boston, MA, USA: Birkhäuser, 1999.
[37] S. Saroiu, P. K. Gummadi, and S. D. Gribble, "A measurement study of peer-to-peer file sharing systems," in *Proc. SPIE/ACM Multimedia Comput. Netw.*, Jan. 2002, vol. 4673, pp. 156–170.
[38] K. C. Sia and J. Cho, "Efficient monitoring algorithm for fast news alerts," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 7, pp. 950–961, Jul. 2007.
[39] K. Sigman and R. Wolff, "A review of regenerative processes," *SIAM Rev.*, vol. 35, no. 2, pp. 269–288, Jun. 1993.
[40] D. Stutzbach and R. Rejaie, "Understanding churn in peer-to-peer networks," in *Proc. ACM IMC*, Oct. 2006, pp. 189–202.
[41] Q. Tan and P. Mitra, "Clustering-based incremental web crawling," *Trans. Inf. Syst.*, no. 4, pp. 1–27, Nov. 2010.
[42] B. Urgaonkar, A. G. Ninan, M. Salimullah, R. Shenoy, and K. Ramamritham, "Maintaining mutual consistency for cached web objects," in *Proc. IEEE ICDCS*, Apr. 2001, pp. 371–380.
[43] "Wikipedia Dumps," 2015 [Online]. Available: http://dumps.wikimedia.org/enwiki/
[44] J. L. Wolf, M. S. Squillante, P. S. Yu, J. Sethuraman, and L. Ozsen, "Optimal crawling strategies for web search engines," in *Proc. WWW*, May 2002, pp. 136–147.
[45] R. W. Wolff, "Poisson arrivals see time averages," *Oper. Res.*, vol. 30, no. 2, pp. 223–231, 1982.
[46] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. Upper Saddle River, NJ, USA: Prentice-Hall, 1989.
[47] M. Yang, H. Wang, L. Lim, and M. Wang, "Optimizing content freshness of relations extracted from the web using keyword search," in *Proc. ACM SIGMOD*, Jun. 2010, pp. 819–830.
[48] H. Yu and A. Vahdat, "Design and evaluation of a continuous consistency model for replicated services," in *Proc. USENIX OSDI*, Jun. 2000, pp. 305–318.



Xiaoyong Li (S'14) received the B.S. degree in computer engineering and M.S. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2005 and 2008, respectively, and is currently pursuing the Ph.D. degree in computer science at Texas A&M University, College Station.

His research interests include peer-to-peer systems, information retrieval, and stochastic analysis of networks.

Daren B. H. Cline received the B.S. degree in mathematics from Harvey Mudd College, Claremont, CA, USA, in 1978, and the M.S. and Ph.D. degrees in statistics from Colorado State University, Fort Collins CO, USA, in 1980 and 1983, respectively.

He joined the Statistics Department at Texas A&M University, College Station, TX, USA, in 1984, and is now a tenured Professor. His current research interests include applied stochastic processes, Markov chain theory, heavy-tailed distributions, and nonlinear time series models.

Dmitri Loguinov (S'99–M'03–SM'08) received the B.S. degree (with honors) in computer science from Moscow State University, Russia, in 1995 and the Ph.D. degree in computer science from the City University of New York, New York, in 2002.

He is currently a Professor in the Department of Computer Science and Engineering at Texas A&M University, College Station. His research interests include P2P networks, information retrieval, congestion control, Internet measurement and modeling.