Multipath Streaming: Fundamental Limits and Efficient Algorithms

Richard Combes
Centrale-Supelec / L2S
richard.combes@supelec.fr

Habib Sidi
INRIA
habisd1@gmail.com

Salah Elayoubi
Orange Labs / SystemX
salaheddine.elayoubi
@orange.com

ABSTRACT

We investigate streaming over multiple links. We provide lower bounds on the starvation probability of any policy and simple, order-optimal policies with matching and tractable upper bounds ([1],[2]).

1. THE MODEL

We consider a file divided in N chunks of unit size, indexed by $n \in \{1,\dots,N\}$. There are $K \geq 1$ links, on which any chunk can be requested. When a requested chunk is received on a link, it is placed in a buffer. After a prebuffering time denoted by $B > 0$ the file is read at unit speed. Namely, at time $n+B$, if chunk n is present in the buffer then it is read, and otherwise starvation occurs. The goal is to design request policies that minimize the starvation probability. We denote by π the request policy where $\pi_n = k$ if chunk n is requested on link k. We assume that if two chunks $n < n'$ are requested on the same link then n is requested before n'. We define $d_k(n) = \sum_{n'=1}^{n} \mathbf{1}\{\pi_{n'} = k\}$, the number of chunks comprised between 1 and n requested on link k. We denote by $X_k(\ell)$ the delay of the ℓ-th chunk requested on link k. Namely, if $\pi_n = k$, chunk n arrives at time $\sum_{\ell=1}^{d_k(n)} X_k(\ell)$. The starvation probability P^N is the probability that there exists a chunk that does not arrive in time. We consider "static" policies where π does not depend on $(X_k(\ell))_{\ell,k}$ and "oracle" policies where π is an arbitrary function of $(X_k(\ell))_{\ell,k}$.

Assumption 1 (i.i.d. delays). *For all k, $(X_k(\ell))_{\ell\geq 0}$ is an i.i.d sequence with expectation μ_k, variance σ_k^2 and cumulant generating function $G_k(a) = \log(\mathbb{E}[e^{aX_k(\ell)}])$. Further, $G_k(a) < +\infty$ on an open neighbourhood of 0.*

[1] An extended version of this work is available at: http://arxiv.org/abs/1602.07112

[2] This work has been performed in the framework of the IDEFIX project, funded by the ANR under contract number ANR-13-INFR-0006

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGMETRICS '16 June 14-18, 2016, Antibes Juan-Les-Pins, France
© 2016 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-4266-7/16/06.
DOI: http://dx.doi.org/10.1145/2896377.2901485

Assumption 2 (Markovian delays). *For all k there exists a continuous time, stationary ergodic Markov chain on a discrete space $\mathcal{S}$ denoted by $(S_k(t))_{t\in\mathbb{R}}$ with stationary distribution m_k and transition rate matrix $Q^k = (q^k(i,j))_{i,j\in\mathcal{S}}$. There exists a function $r : \mathcal{S} \to \mathbb{R}^+$ such that for all ℓ, k:*

$$\tau_k(\ell) = \min\left\{t \geq 0 : \int_0^t r(S_k(u))du \geq \ell\right\},$$

with $X_k(\ell) = \tau_k(\ell) - \tau_k(\ell-1)$ and $\mu_k = \mathbb{E}[X_k(\ell)]$.

We define $r_k = 1/\mu_k$, the average data rate of link k, and $R = \sum_{k=1}^{K} r_k$ the sum of data rates. We distinguish three regimes: underload ($R > 1$), critical ($R = 1$) and overload ($R < 1$). We define the frequency vector $f = (f_1, ..., f_K)$, with $f_k = r_k/R$. We denote by $P^N(\pi, B)$ the starvation probability for N chunks, prebuffering time B and policy π.

2. PERFORMANCE LIMITS

Theorem 1 is a lower bound on the starvation probability that holds for all oracle policies. For large files ($N \to \infty$), there are sharp transitions between regimes: to ensure that $P \not\to 1$ we require $B = \mathcal{O}(1)$ (underload), $B = \mathcal{O}(\sqrt{N})$ (critical) and $B = \mathcal{O}(N)$ (overload).

Theorem 1. *The following holds for all oracle policies π.* *(i) For all $B \geq 0$ and $N \geq 1$ we have:*

$$P^N(\pi, B) \geq \mathbb{P}\left[\exists n \in \{1,...,N\} : \sum_{k=1}^{K} D_k(n,B) < n\right]$$

$$D_k(n,B) = \max\{d \geq 0 : \sum_{\ell=1}^{d} X_k(\ell) \leq B + n\}.$$

(ii) Consider i.i.d. delays. If $R \leq 1$, for all $b \geq 0$ we have:

$$\liminf_{N\to\infty} P^N(\pi, (R^{-1}-1)N + b\sqrt{N}) \geq \prod_{k=1}^{K} \Psi\left(\frac{b}{\sigma_k\sqrt{f_k}}\right),$$

with $\Psi(x) = \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} e^{-\frac{z^2}{2}}dz$.

3. EFFICIENT ALGORITHM

To obtain an efficient policy, chunks should be requested on link k at frequency f_k, so that $d_k(n) \approx nf_k, \forall k, n$. Policy π is f-upper balanced if $d_k^\pi(n) \leq (n+K-1)f_k$ $\forall k, n$.

Proposition 1. *Consider π such that for all $n \geq 0$:*

$$\pi_n \in \arg\min_k \frac{d_k^\pi(n-1)+1}{f_k}$$

with ties broken arbitrarily. Then π is f-upper balanced.

4. PERFORMANCE: I.I.D. DELAYS

Consider i.i.d. delays and define $F_k(a) = G_k(a) - a/f_k$. Define $a_k^\star = +\infty$ if $X_k(\ell) < 1/f_k$ a.s. and $a_k^\star = \max\{a \geq 0 : F_k(a) = 0\}$ otherwise. We calculate $a_k^\star$ below for exponential delays and sub-Gaussian delays, which includes bounded and Gaussian delays. We say that $X_k(\ell)$ is v_k^2-sub-Gaussian if $G_k(a) \leq a\mu_k + a^2 v_k^2/2, \forall a \geq 0$.

Proposition 2. *Consider $R > 1$.*
(i) If $X_k(\ell) \sim Exp(r_k)$, then $a_k^\star = r_k(1 + W(-Re^{-R})/R)$ with W the Lambert function.
(ii) If $X_k(\ell)$ is v_k^2-sub-Gaussian then $a_k^\star \geq 2\mu_k(R-1)/v_k^2$.

Theorem 2 gives upper bounds on the starvation probability of upper balanced policies, and shows that they are order optimal: the pre-buffering times have the same scaling as the lower bound of Theorem 1.

Theorem 2. *Let π be f-upper balanced. Define $N_k = f_k N$ and $\mathcal{G}_k = \{a : F_k(a) \geq 0\}$.*
(i) For all $b \geq 0$ and $N \geq 1$ we have:

$$P^N(\pi, b+K-1) \leq 1 - \prod_{k=1}^{K}\left[1 - \min_{a_k \in \mathcal{G}_k} e^{N_k F_k(a_k) - a_k b}\right].$$

(ii) If $R > 1$, we have $a_k^\star > 0$ and for all $N \geq 1$ and $b \geq 0$:

$$P^N(\pi, b+K-1) \leq 1 - \prod_{k=1}^{K}\left[1 - e^{-a_k^\star b}\right].$$

(iii) If $R \leq 1$, and $X_k(\ell)$ is v_k^2-sub-Gaussian for all k, then for all $b \geq 0$ and $N \geq 1$:

$$P^N(\pi, (R^{-1}-1)N + b + K - 1) \leq 1 - \prod_{k=1}^{K}\left[1 - e^{-\frac{b^2}{2v_k^2}}\right].$$

5. PERFORMANCE: MARKOV DELAYS

For Markovian delays, the problem is mostly intractable, and we consider a regime where $S_k(t)$ evolves on a "faster time scale" than the streaming flow of interest. This regime makes sense since typical streaming flows are long while link variability is caused by short phenomena such as fading, medium access protocols and short-lived elastic flows. We replace $(S_k(t))_t$ by the accelerated process $(S_k(\phi t))_t$ with speed $\phi > 0$. We use Lemma 1, which shows that the amount of data received on a link can be approximated by a Wiener process. We identify $r(.)$ with $(r(i))_{i \in \mathcal{S}}$.

Lemma 1 (Bhattacharya, 82). *Define $G^\phi(t) = \sqrt{\phi}\int_0^t (r(S_k(\phi u)) - r_k)du$. Consider $g^k = (g^k(i))_{i\in\mathcal{S}}$ a solution to the Poisson equation: $Q^k g^k = r(.) - r_k$. Define the asymptotic variance $\overline{\sigma}_k^2 = -2\sum_{i\in\mathcal{S}} r(i) g^k(i) m_k(i)$. Then $G^\phi(.)$ converges to a Wiener process with drift 0 and variance $\overline{\sigma}_k^2$, when $\phi \to \infty$.*

Theorem 3 gives upper bounds on the starvation probability of upper balanced policies. Statement (i) shows that for $R > 1$ fixed and $\phi \to \infty$ we have $P \to 0$ i.e. the link variability disappears due to the ergodic theorem. Statement (ii) deals with the case where R depends on ϕ and approaches 1 as $\phi \to \infty$. In cases of interest $\overline{\sigma}_k^2$ can be calculated explicitly making our performance bounds tractable as shown below.

Theorem 3. *Let π be f-upper balanced.*
(i) Consider $R > 1$ and $b > 0$ fixed. Then for all $N \geq 1$:

$$P^{N,\phi}(\pi, b+K-1) \underset{\phi\to\infty}{\to} 0.$$

(ii) Consider $C_1, C_2 \geq 0$, define $b^\phi = C_2/\sqrt{\phi}$ and assume that $R = R^\phi \equiv 1/(1 - C_1/\sqrt{\phi})$. Then for all $N \geq 1$:

$$\limsup_{\phi\to\infty} P^{N,\phi}(\pi, b^\phi + K - 1)$$
$$\leq 1 - \prod_{k=1}^{K}\left(1 - \mathbb{P}\left[\max_{t\in\{K,\dots,N+K-1\}} \overline{\sigma}_k W(t) - r_k C_1 t \geq f_k C_2\right]\right).$$

with $(W(t))_t$ a standard Wiener process.
(iii)(a) If $C_1 = 0$, $R^\phi = 1$ and for all $N \geq 1$:

$$\limsup_{\phi\to\infty} P^{N,\phi}(\pi, b^\phi + K - 1)$$
$$\leq 1 - \prod_{k=1}^{K}\left(1 - 2\Psi\left(\frac{f_k C_2}{\overline{\sigma}_k\sqrt{N+K-1}}\right)\right)$$

(iii)(b) If $C_1 > 0$, for all $N \geq 1$:

$$\limsup_{\phi\to\infty} P^{N,\phi}(\pi, b^\phi + K - 1) \leq 1 - \prod_{k=1}^{K}\left(1 - e^{-2\frac{r_k^2 C_1 C_2}{\overline{\sigma}_k^2}}\right).$$

6. SOME RELEVANT LINK MODELS

6.1 Wireless links with random access

We consider Bianchi's model with one back-off stage, window size W and collision probability p (calculated through a fixed point equation). If a transmission is successful one transmits a frame, otherwise one waits for a duration uniformly distributed in $[0, W]$. A chunk is composed of n_f frames. The cumulant generating function of delays is:

$$G(a) = n_f\left(a + \log\left(\frac{p}{1 - (1-p)h(aW)}\right)\right)$$

with $h(a) = (e^a - 1)/a$ and a such that $(1-p)h(aW) < 1$.

6.2 Wireless ON-OFF channels

A link is shared between a secondary user and a primary user whose activity follows a two-states Markov process independent from the secondary user activity. The secondary user transmits at rate 1 when the primary user is not active. The transition rate matrix is $Q = \begin{pmatrix} -\beta & \beta \\ \alpha & -\alpha \end{pmatrix}$. The stationary distribution is $m = (\frac{\alpha}{\alpha+\beta}, \frac{\beta}{\alpha+\beta})$, the expected data rate is $r = \frac{\beta}{\alpha+\beta}$. The Poisson equation yields: $\overline{\sigma}^2 = \frac{2\beta\alpha}{(\alpha+\beta)^3}$.

6.3 Links with short lived flows

A link is shared between the streaming flow and $S(t)$ short flows following an M/M/1 process with load $\rho < 1$. When there are n small flows, the streaming flow transmits at rate $r(n)$ (e.g. $r(n) = 1/(1+n)$ for fair rate sharing). The stationary distribution is $m(n) = \rho^n(1-\rho)$. The expected data rate is $\overline{r} = \sum_{n\geq 0} r(n)\rho^n(1-\rho)$. Define $\overline{R}(n) = r(n) - \overline{r}$. Solving the Poisson equation we get:

$$\overline{\sigma}^2 = 2\rho\sum_{n\geq 0}\sum_{i=0}^{n-1}\overline{R}(n)\overline{R}(i)(\rho^n - \rho^i).$$