

Understanding latency in Kaltura live streams

Last Modified on 11/20/2025 2:26 pm IST

 This article is designated for administrators.

About

This article explains what latency is in Kaltura live streams, what affects it, how segment duration impacts delay, and how to enable low-latency delivery options.

What is latency?

Latency is the delay between when the camera captures video and when it's displayed in the player. Several factors contribute to the overall delay:

- Local encoder performance
- Network uplink (encoder to Kaltura cloud)
- Media processing (cloud transcoding, packaging, encryption if needed)
- Network downlink (Kaltura cloud to CDN to player)
- Player internal buffers

A significant contributor to the overall latency is the player behavior. This is an inherent behavior to any HTTP-based adaptive streaming, where the video is segmented into small chunks. The player downloads several segments (depending on the player configuration) before playback starts to maintain a safe buffer and avoid interruptions.

For example, the iOS native player downloads three 10-second segments before playback, adding about 30 seconds of delay.

Use cases that require lower latency

For many live streams, latency doesn't impact the viewer experience. Latency becomes important when timing or interaction are part of the workflow, such as:

- Webcasting with Q&A or polls, and other interactive sessions that require two-way engagement
- Betting or bidding
- Second-screen experiences
- Sports
- Real-time conversation or chat-style experiences

Testing results across devices and browsers

Kaltura regularly tests streaming latency across a wide range of desktop and mobile devices and browsers. These tests show that shorter segment durations consistently reduce end-to-end delay:

- **2-second segments:** typically **10-15 seconds** latency
- **4-second segments:** typically **18-23 seconds** latency
- **6-second segments (default):** typically **25-30 seconds** latency

These ranges reflect results from Kaltura's internal testing and may vary slightly based on network quality, device capabilities, and player buffering.

How to reduce latency

There is always a trade-off between lowering latency and maintaining a stable, buffer-free experience. Shorter segment durations can reduce latency, but they may increase sensitivity to unstable networks. For this reason, shorter segments are not enabled by default.

If lower latency is required, your Kaltura representative can enable and configure **Dynamic Segment Duration** for your account. This feature allows Kaltura to apply shorter segment durations, which reduce end-to-end latency.

Kaltura also supports a **low latency** streaming mode (currently in Beta). For more information, see our article [Low latency streaming guidelines](#).

Create Live Stream Entry

Stream Type	Kaltura Live Streaming (HDS/HLS/DASH) ▾
Name *	<input type="text"/>
Description	<input type="text"/>
Live Transcoding Profiles	Passthrough ▾
Live DVR	<input type="checkbox"/> Off <small>Seeking to previous parts of the live stream. The DVR window is 2 hours by default.</small>
Recording	<input checked="" type="checkbox"/> On <div>Append recorded content to a single entry ▾</div>
Preview Mode	<input type="checkbox"/> Off
Low Latency (Beta)	<input checked="" type="checkbox"/> On ⓘ

Create



Low-latency delivery is more sensitive to unstable networks. Enable this mode

when viewers have consistent bandwidth or when low delay is essential for the experience.
