



# VIDEO STREAMING LATENCY REPORT 2021



wowza.com

# 2021 Video Streaming Latency Report

Low-latency video delivery has always been the holy grail of streaming. But the need to replicate real-life exchanges on a global scale has never been so crucial as it has in the past two years.

It started in early 2020 [when everything went virtual](#). Conferences, shopping centers, casinos, and doctors' offices became digital places accessed via the internet. From there, many in-person experiences continued to transform into hybrid environments connecting both remote and on-site participants.

Content distributors in every industry were forced to adapt. Manufacturers looked to [tele-welding](#). Commercial real estate firms began leveraging [live 360° streaming](#). And [live event producers experimented](#) with every option for increasing engagement via interactivity. For many, low-latency streaming went from a nice-to-have to a key requirement.

This report gathers data from more than 200 broadcasters across the globe, in industries ranging from media and entertainment to fitness, surveillance, and everything in between. We examine the impact these changes have had on the industry and the state of video streaming latency in 2021.

For everyone who participated, we are forever grateful. It's clear that many content distributors now have their sights set on [real-time streaming technologies](#) — and we're excited to share how Wowza has been working to deliver just that.

## Report Highlights

- **Interactivity is the name of the game.** Popular motivators for decreasing latency include audience polling, Q&A, bidding, and the like.
- **Use cases keep growing.** While live sports broadcasts still lead the way, 16% of those surveyed streamed content in categories ranging from underwater exploration to construction safety.
- **Most live streams aren't live.** Or, to put it another way, more than 52% of survey participants experience more than a three-second lag in their live broadcasts.
- **RTMP and HLS maintain their stronghold, but next-generation formats are on the rise.** These include WebRTC, SRT, Low-Latency HLS, and low-latency CMAF for DASH.
- **The industry needs real-time streaming at scale.** The majority of broadcasters surveyed are aiming to achieve real-time delivery to audiences of 300+ viewers (*and luckily, Wowza has a [solution](#)*)

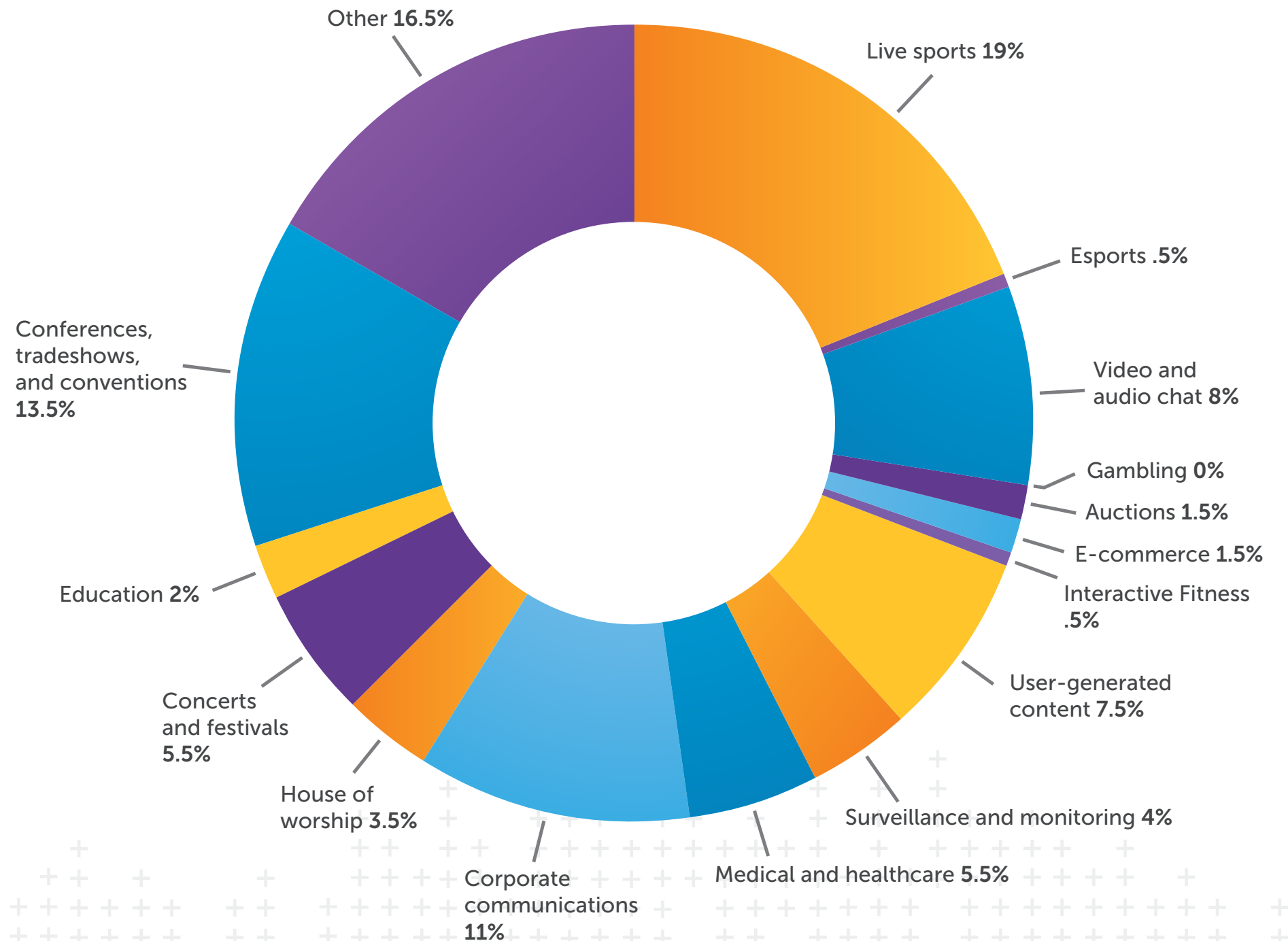
# Q1 What type of content are you streaming?

## USE CASE

To repeat an assertion made in our 2019 report: The live streaming landscape is more diverse than ever. Industries ranging from [healthcare](#) to [e-commerce](#) are digitally transforming at an unprecedented rate by leveraging video technology.

Live sporting events remain the leading type of broadcast, and the demand for low latency keeps growing. After all, nobody wants to learn the final score on their Twitter feed while the stream they're watching lags behind.

Conferences, tradeshows, and conventions have also become major players. For this, we have COVID-19 to thank, which forced traditionally in-person [events](#) to adopt virtual and hybrid models. Interactive technologies are fundamental to bringing these broadcasts to life, which is why speedy video delivery remains a top priority.

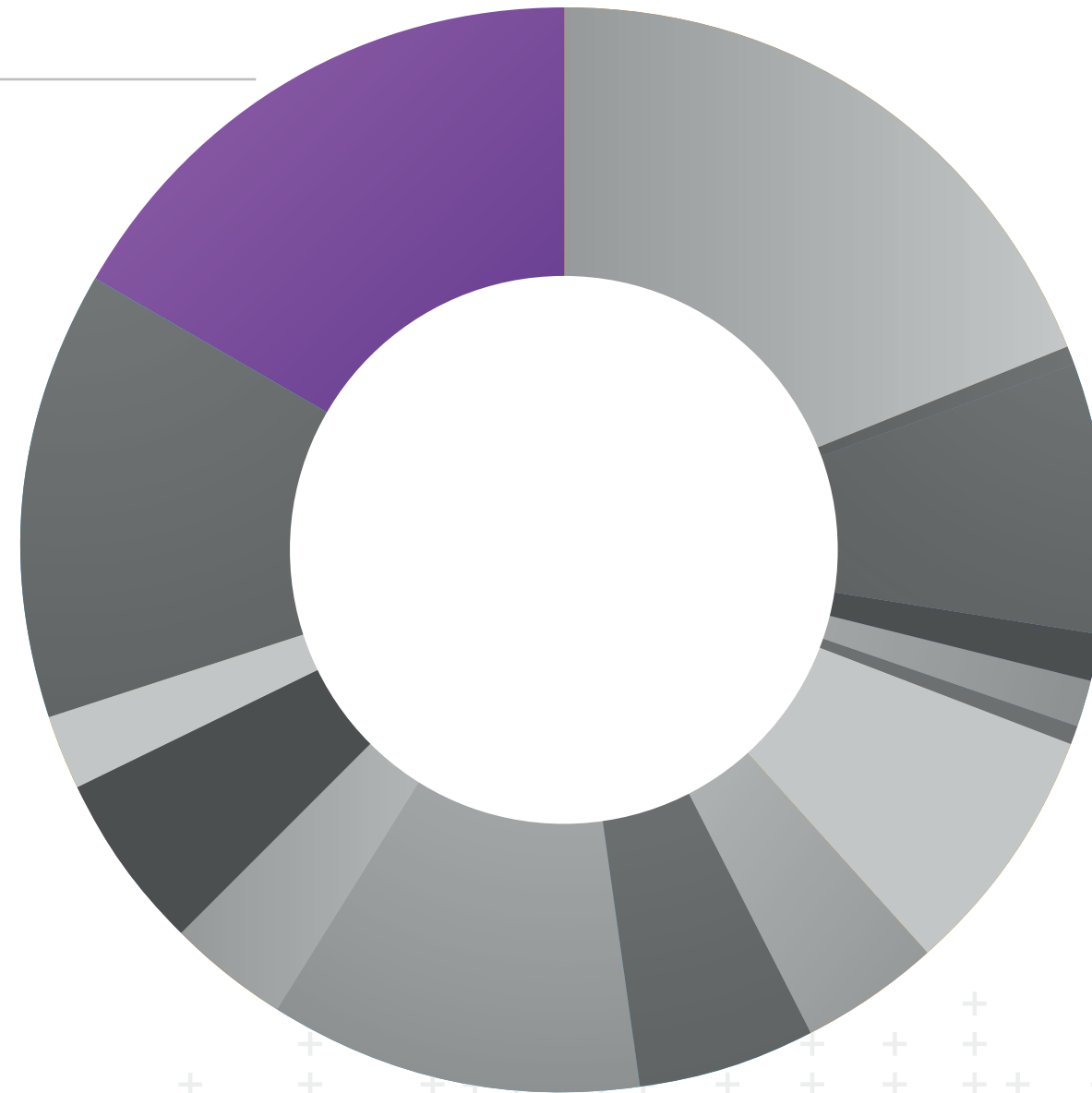


# Q1 What type of content are you streaming?

## USE CASES CONTINUE TO EXPLODE

### Other (Please Specify)

- Audiobooks
- 360° streaming and VR/AR
- Local government meetings
- News
- Underwater exploration
- Podcasts
- Dance recitals
- Funerals
- Construction Safety

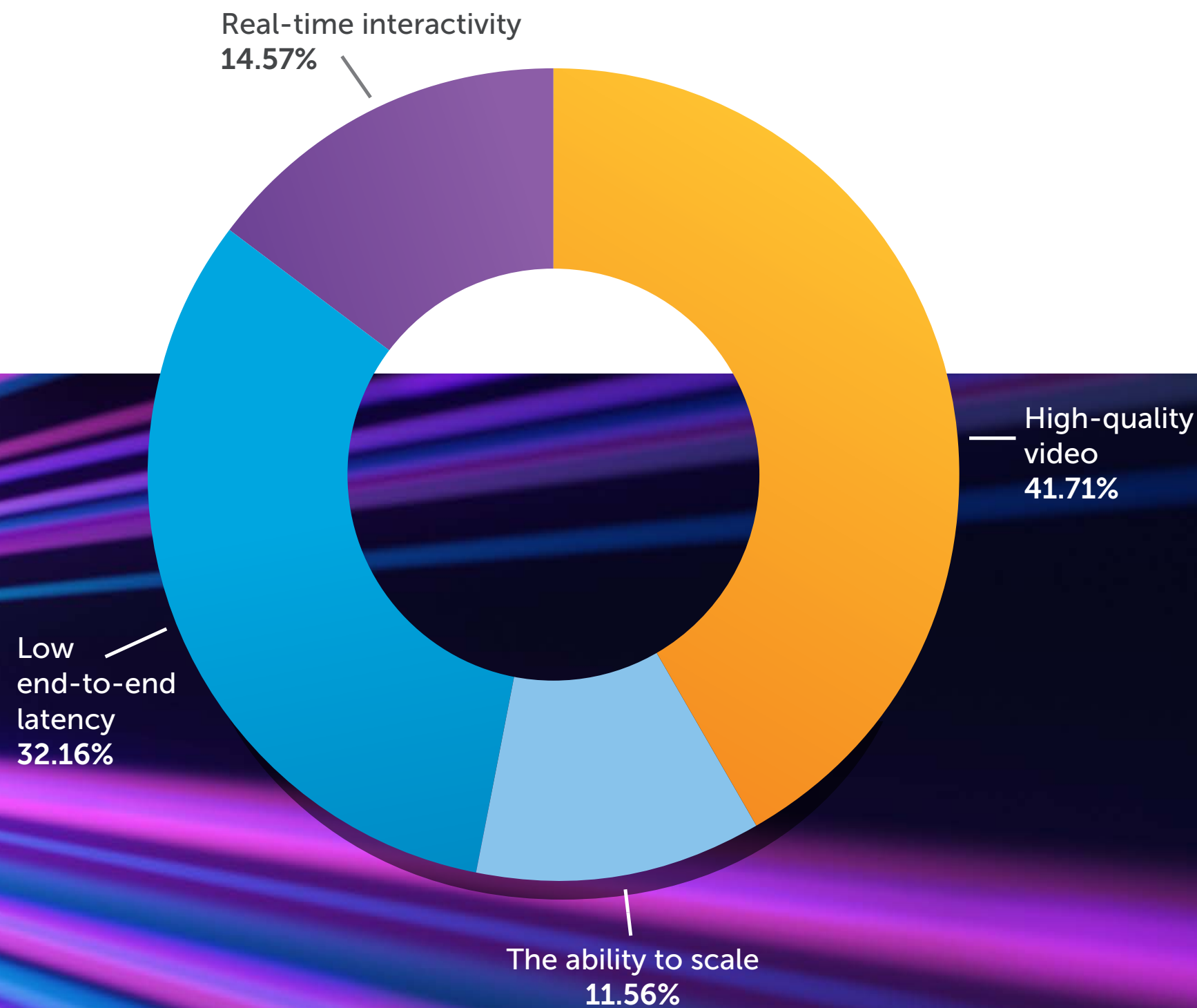


Next to live sports, 16.5% of respondents fell in the 'other' category. Several indicated that they stream a combination of content types – but a handful of unusual use cases also popped up, as listed above.

# Q2 What is the most important UX factor for your use

## CRITICAL CAPABILITIES

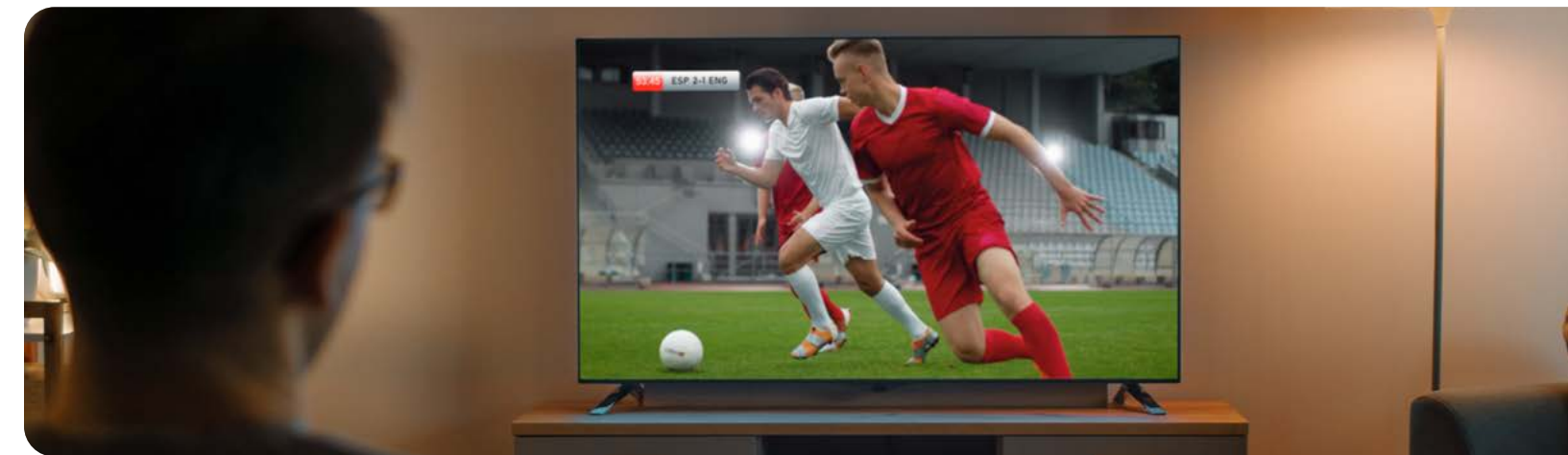
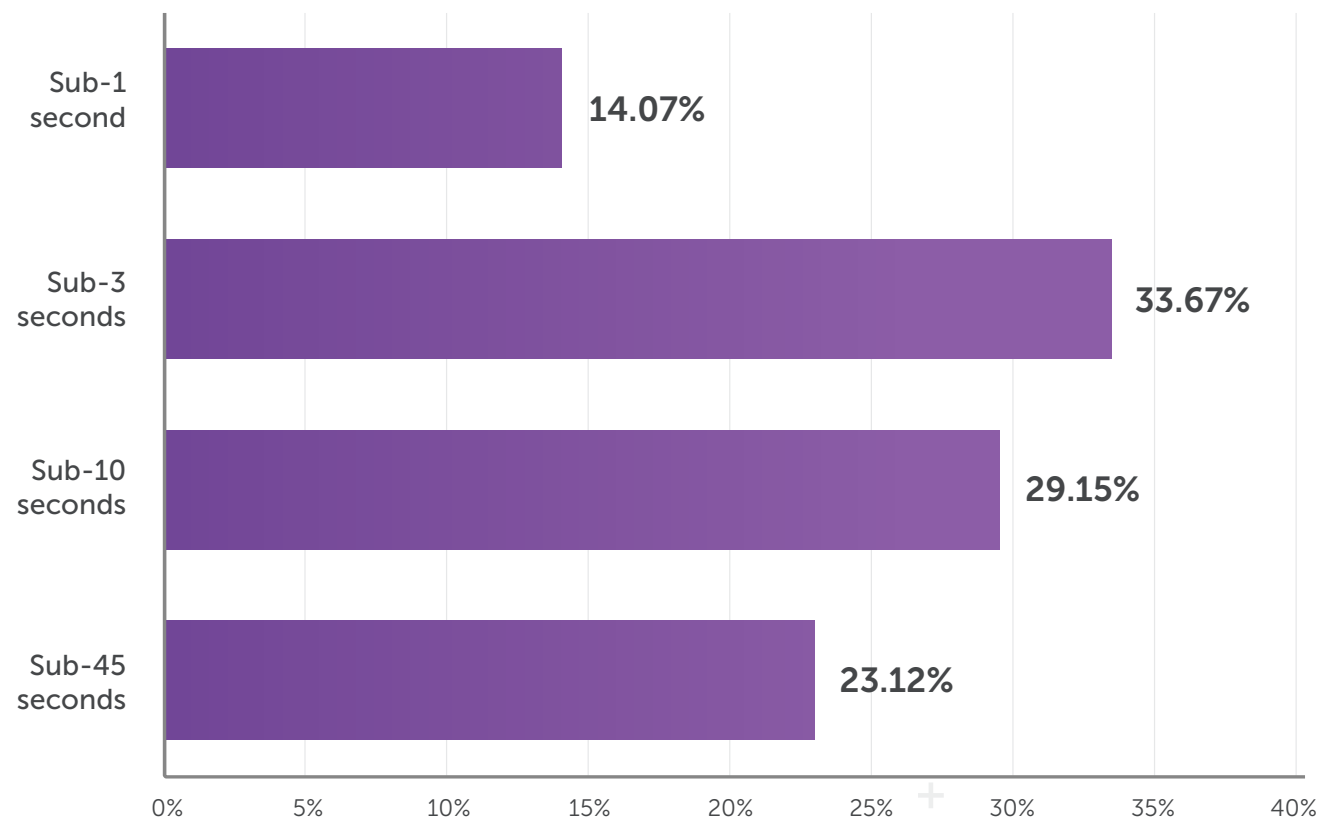
High-quality streaming is a must, but low-latency delivery is hot on its heels. And because real-time interactivity requires low latency, it stands to reason that minimizing delay is the most critical capability for many of today's live broadcasters.



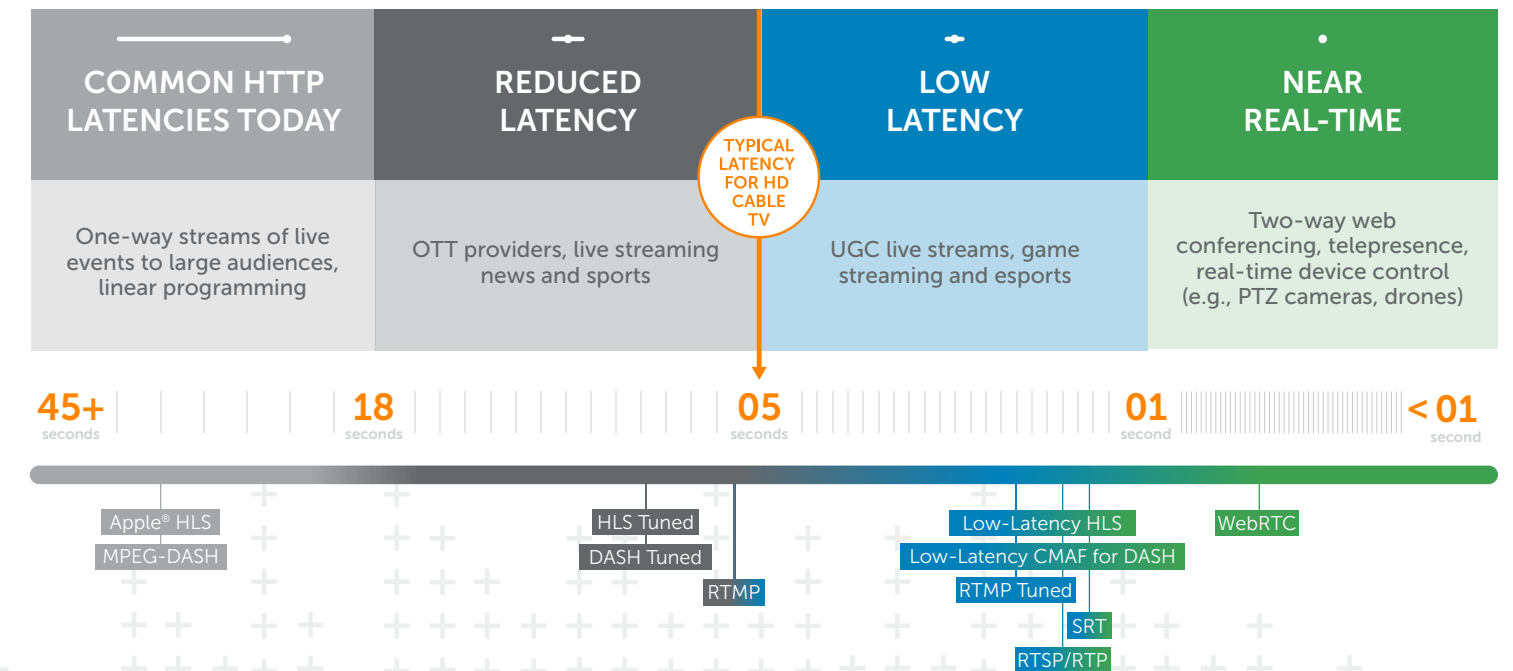
# Q3 How much latency are you currently experiencing?

## CURRENT LATENCY

Don't be fooled by this graph. Although sub-three-seconds stands out as the most prominent category, more than 52% of survey participants are stuck with a delay north of that. And because standard TV often clocks in at the five-second mark, that means a good chunk of OTT broadcasters still lag behind [traditional cable and satellite services](#).



## STREAMING LATENCY AND INTERACTIVITY CONTINUUM



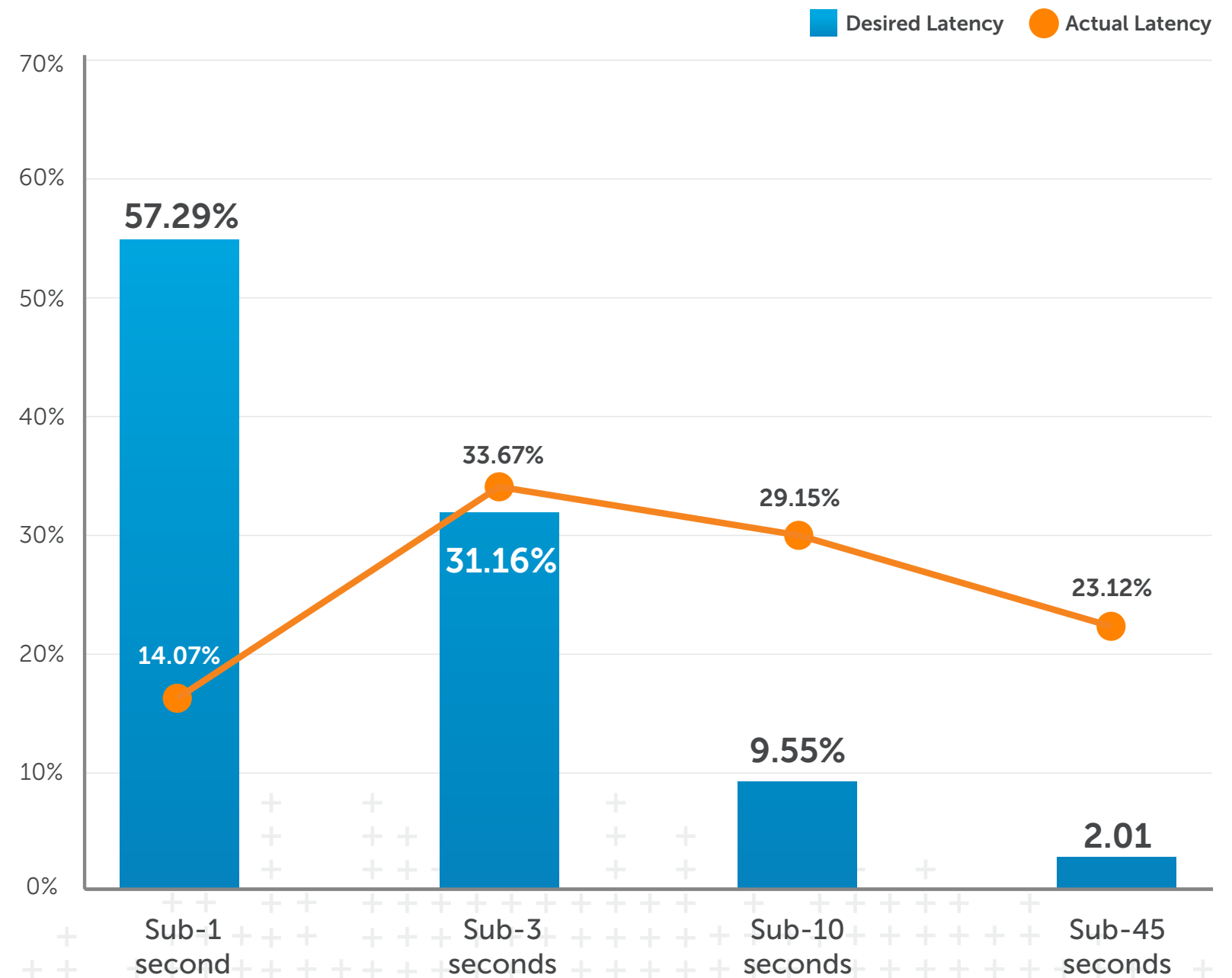
# Q4 How much latency do you hope to achieve in the future?

## DESIRED LATENCY

It should come as no surprise that broadcasters across every industry want to drive latency down – with more than 57% hoping to achieve sub-one-second delivery. That said, although real-time delivery sounds like a good thing for anyone to strive for, we’d recommend being open to some lag when your use case allows.

Why? Because configuring streams for speedy delivery can introduce complexity and costs that aren’t necessary. Most passive broadcasts should be safe in the sub-ten-second range, and it’ll allow for a smoother, more reliable stream.

On the flip side, for content distributors who simply can’t let the seconds pile up – such as those deploying [interactive experiences](#) or [mission-critical applications](#) – real-time delivery becomes non-negotiable. This is easier said than done, though, as indicated by the disparity between current and desired latencies in this graph.



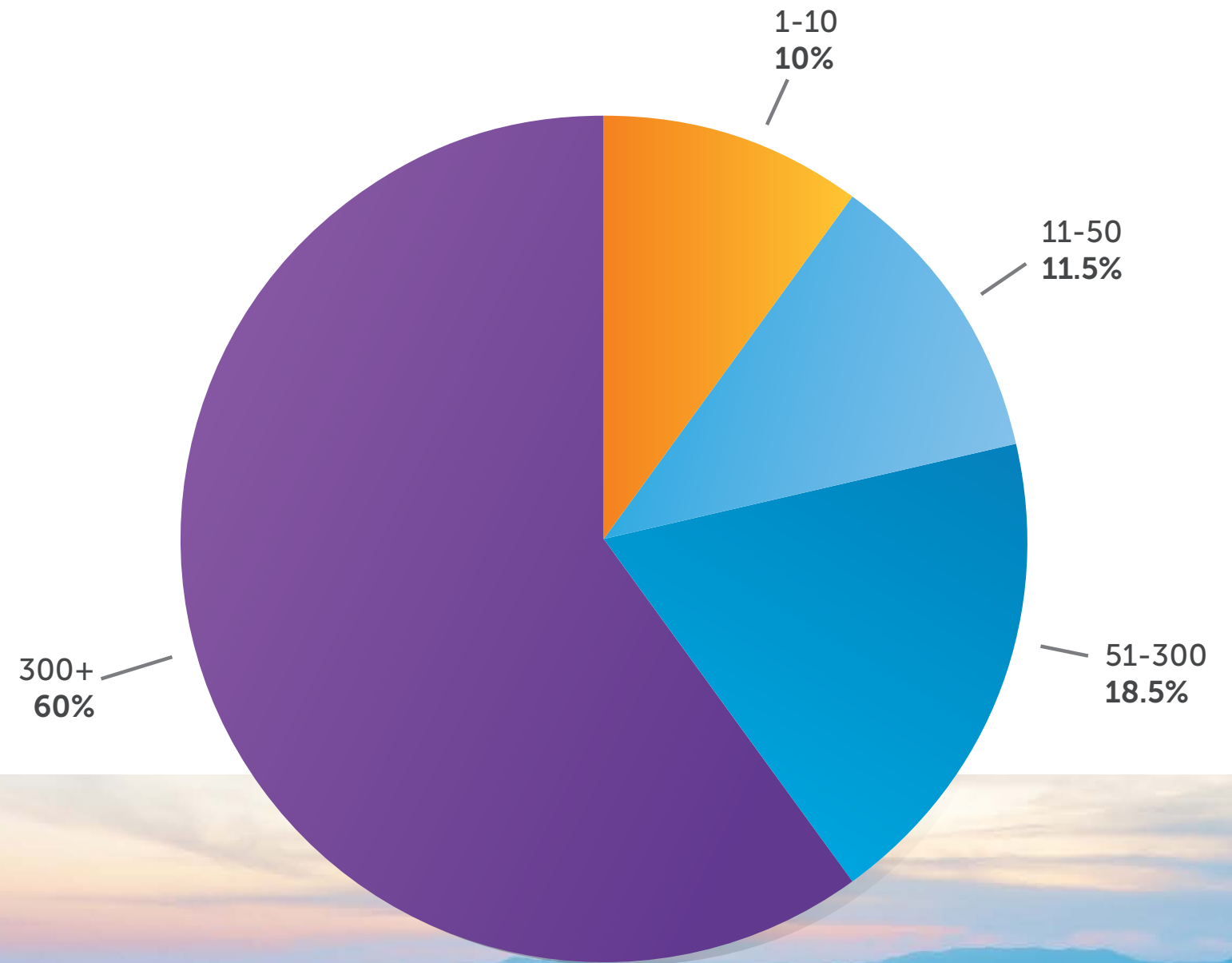
# Q5 What is your audience size?

## SCALE

Combining the data from this question with that of the previous one leads to one conclusion: The majority of broadcasters surveyed aim to achieve real-time delivery to audiences of 300+ viewers.

These two requirements (low latency and scale) used to be at odds with one another. That's because the only format that ensures sub-one-second streaming is [Web Real-Time Communications \(WebRTC\)](#), which was designed for small video chat environments rather than one-to-many broadcasts.

Luckily, though, our new [Real-Time Streaming at Scale](#) feature for [Wowza Video](#) brings you the best of both worlds, making this goal within reach.





# Q6 Are you currently using low-latency streaming services?

## LOW-LATENCY SERVICE USE

Most broadcasters still haven't made the switch to low-latency services. Why? Lacking support, stability, and budget all come into play — as well as the fact that, for some, minimizing delay just isn't necessary.

Here's how one respondent described it:

"[Latency] is not important. Who cares if the stream is delayed a few seconds? Heck, even a few minutes. It's not like our viewers would even know."

### Top reasons for not implementing low-latency services:

- Don't need it
- Still testing
- Don't have enough information
- Lack of vendor support
- Too costly
- Too expensive or difficult to scale
- Challenges with reliability and stability
- Need additional development resources
- Waiting for low-latency formats to mature and finalize



# Q7 What problems would real-time streaming solve for you?

## NEED FOR REAL-TIME STREAMING

"Conversations become more authentic." **"Smoother experience."**

"Replays, scoreboards, commentary." "Make clients in the military sector happy."

**"Interactivity with online audience."**

"Mobile users can place timely bids." **"Polling, surveys, Q&A."**

"Bring parity between remote and on-site audiences for hybrid events." "Ability to compete in new markets."

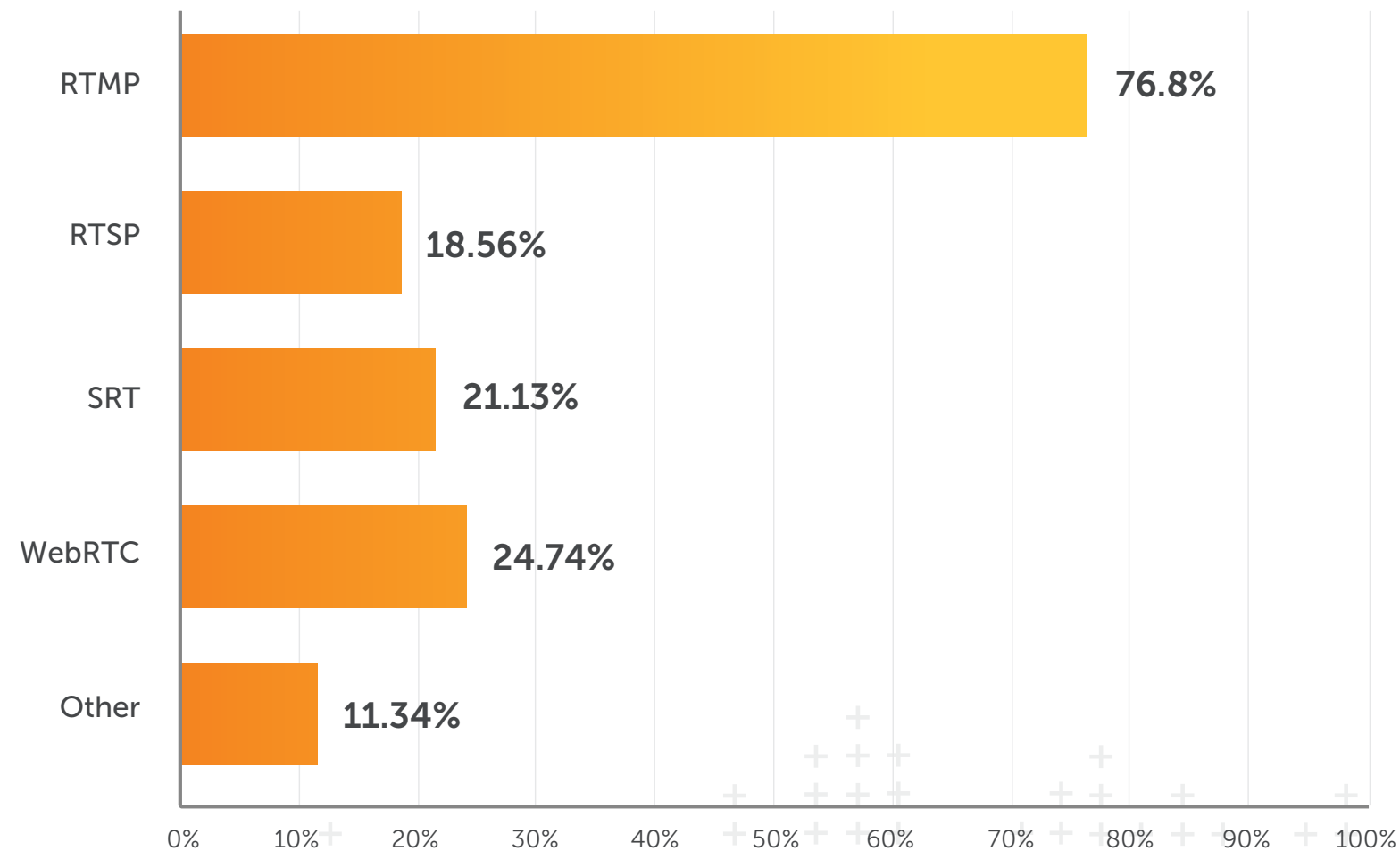
**"Match the timing of linear broadcasts."**

# Q8 Which streaming formats are you currently using for ingest?

## INGEST FORMATS

[The Real-Time Messaging Protocol \(RTMP\)](#) remains the number one format for [ingest](#), with more than 76% of broadcasters indicating they use it. We can chalk up RTMP's popularity to the fact that it's widely supported on the ingest side and RTMP-based workflows are well defined. But that's not to say it'll remain this way.

WebRTC comes in second, and we expect usage to keep growing. For one, it's the fastest technology of the bunch and can also be used from end to end. Adoption of WebRTC on the delivery side is also picking up, which is reflected on the next page.

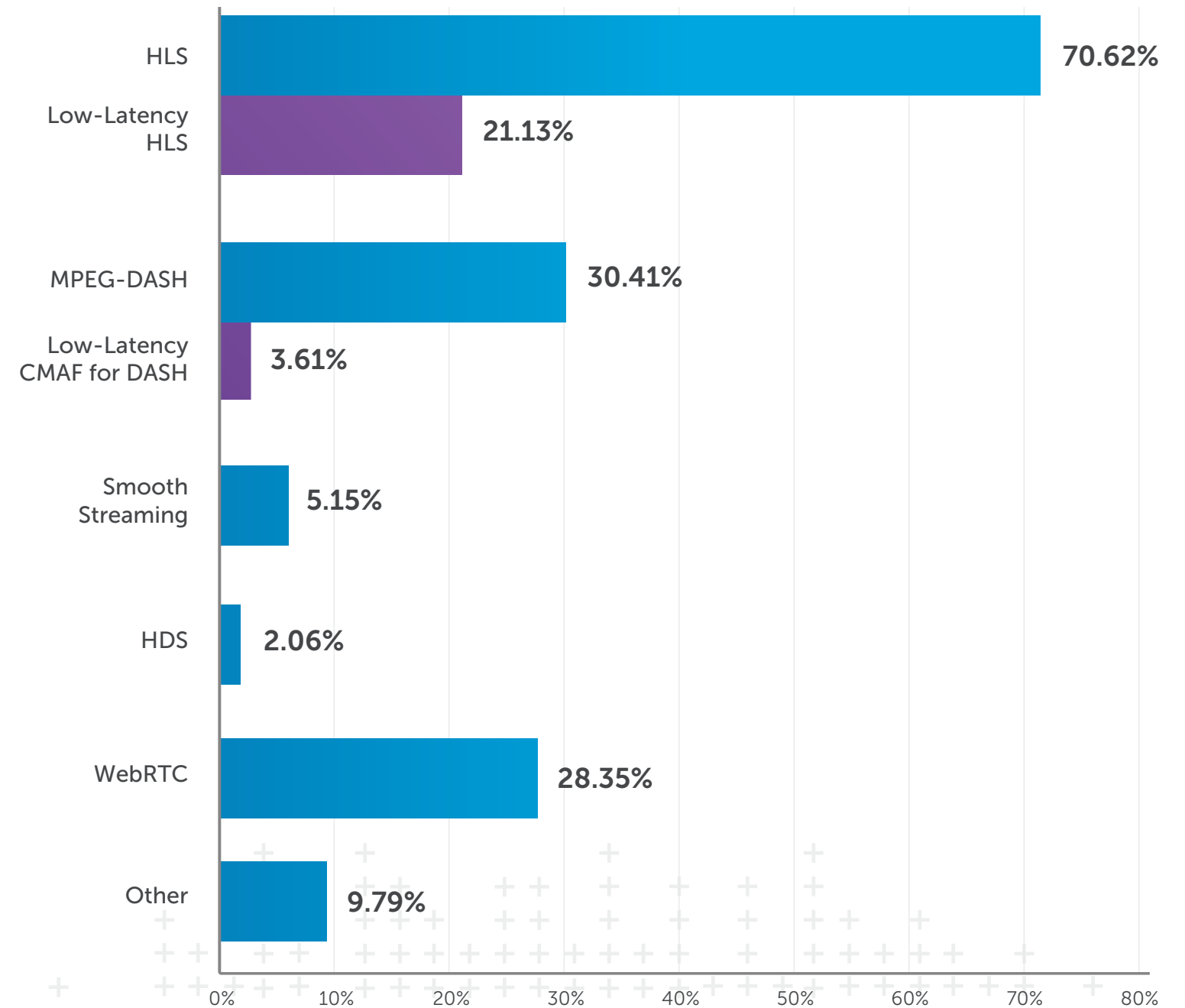


# Q9 Which streaming formats are you currently using for delivery?

## DELIVERY FORMATS

Predictably, more than 70% of respondents deliver their live streams using [Apple's HTTP Live Streaming \(HLS\) protocol](#). From there, [MPEG-DASH](#) and WebRTC are neck in neck. Smooth Streaming and HDS are dying a slow death, with vendor support waning.

It's worth pointing out that the past two graphs have added up to more than 100%. That's because multi-protocol delivery and hybrid workflows are becoming the norm. Many respondents indicated using a handful of formats on each end of the streaming workflow, with RTMP in and HLS popping up most often.



**Funny story:** When we initially launched this survey, we got attacked by bots. The first indicator that the data coming in was suspect? Smooth Streaming was leading the pack in delivery formats. Once we scrubbed the data and added bot-blocking measures, the more accurate picture depicted here revealed itself.

# Q10 Where do you experience the largest delay in your workflow?

## SOURCE OF DELAY

Being as HLS came in as the number one delivery format, those who marked player buffer as their most significant source of delay were right on the money. But for the 23% of respondents who weren't sure where their delay stems from, let's look at where latency creeps in from capture to playback.

### Encoding:

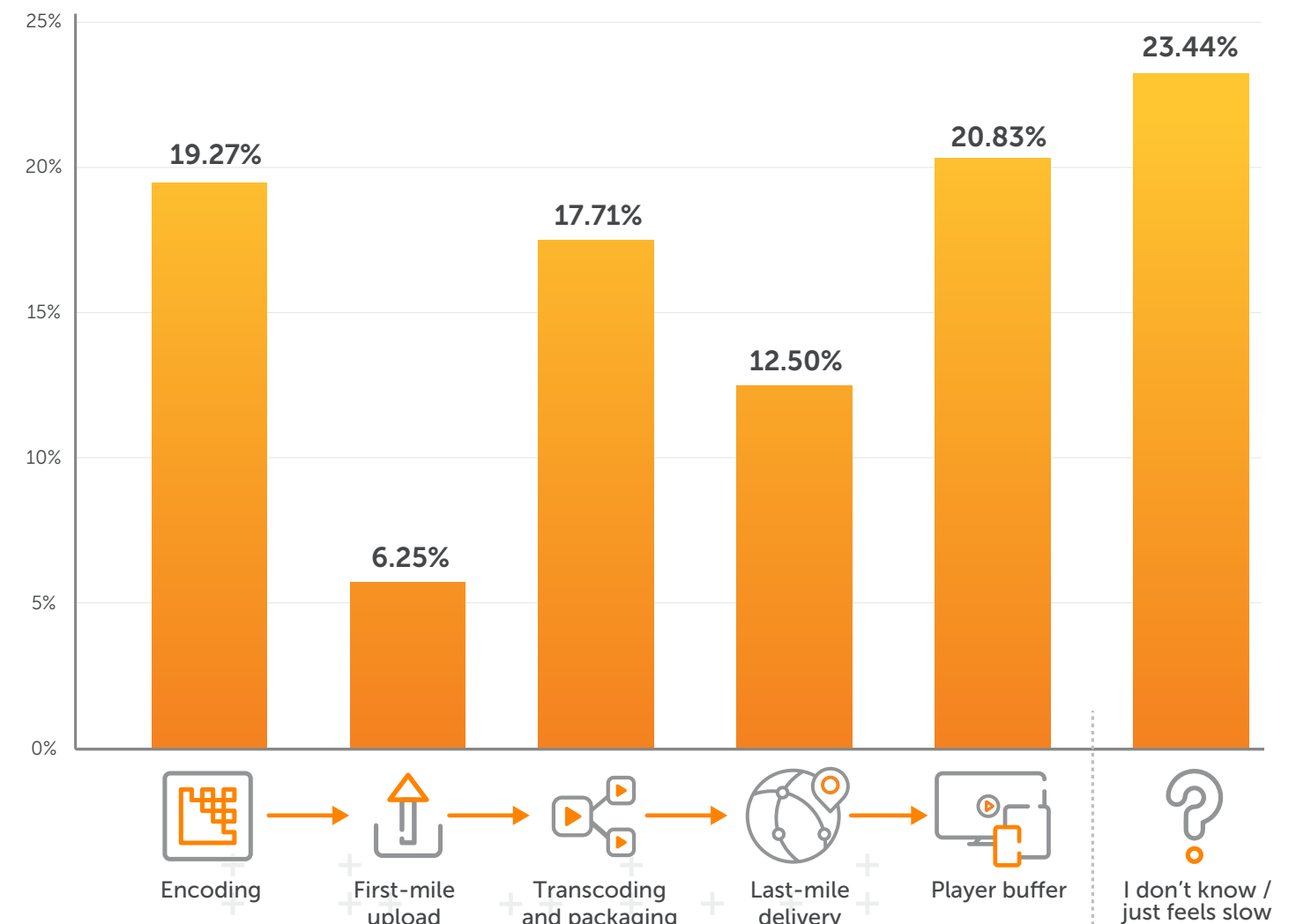
Bitrate, resolution, which [codec](#) you use, and even segment size impact the speed of [video encoding](#). The higher the bitrate and resolution, the longer encoding will take. It's a good idea to use a video transcoding solution like Wowza to ensure efficiency at the encoder.

### First-Mile Upload:

Contribution delays often result from transmitting data over suboptimal networks and synchronizing multiple video sources. By choosing a protocol designed for low-latency content acquisition like [SRT](#), this can easily be avoided in remote locations. Otherwise, the connection type is key.

### Transcoding and Packaging:

Traditional streaming protocols such as RTSP and RTMP support low-latency streaming. But they aren't supported by many players. Many broadcasters choose to transport live streams to their media server using RTMP and then [transcode](#) it for multi-device delivery. The process itself injects latency, as do common delivery protocols like HLS.



# Q10<sup>Cont.</sup> Where do you experience the largest delay in your workflow?

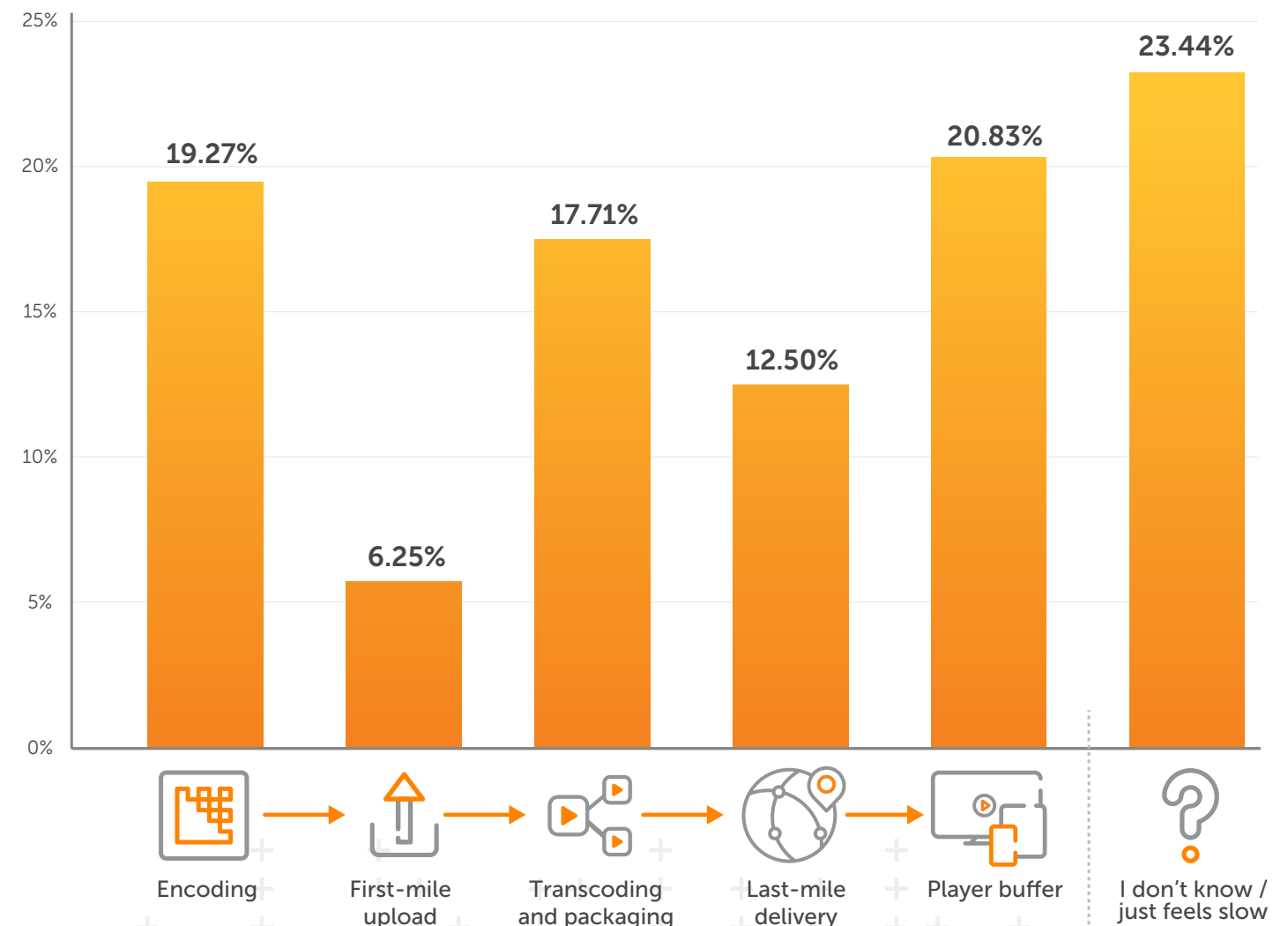
## SOURCE OF DELAY

### Last Mile:

The farther your viewers are from the edge server, the longer it'll take to distribute a stream. This part of the workflow is largely outside of your control. End-users' proximity to the [CDN](#) edge and their network conditions will influence last-mile delivery. That said, some protocols like WebRTC weren't designed for large-scale delivery via a typical CDN, meaning you'll need a custom solution like our [Real-Time Streaming at Scale feature](#) to keep this step up to speed.

### Player Buffer:

Many specifications require a certain number of segments to be loaded before playback can begin. This buffer is intended to improve the viewer experience. When real-time delivery is essential, you'll want to swap out a traditional HTTP-based format for WebRTC.

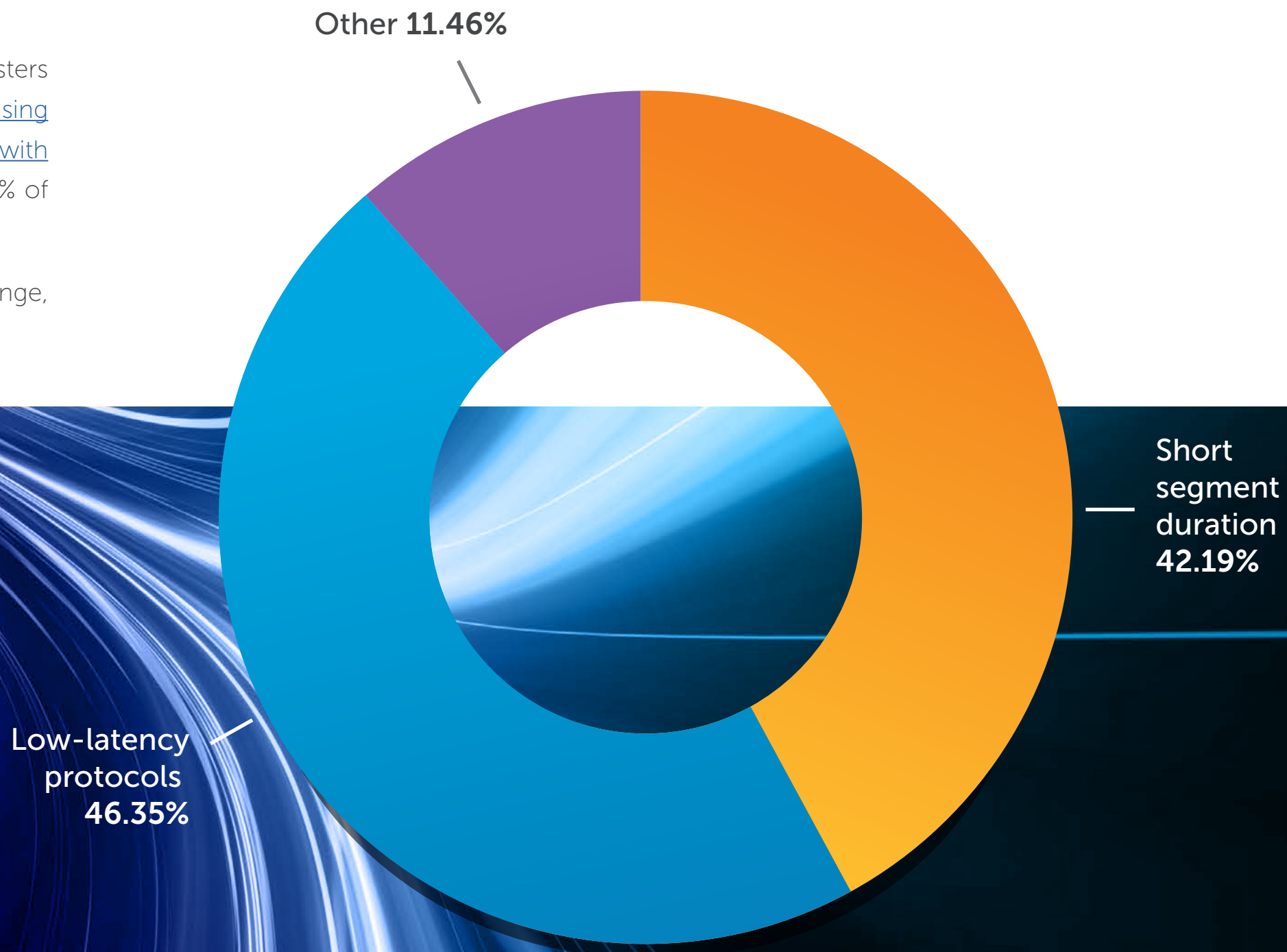


# Q11 How are you currently reducing latency?

## LOW-LATENCY TACTICS

As latency requirements become more aggressive, broadcasters are transitioning from the tried-and-true approach of [decreasing segment length](#) (42%) to implementing [technologies designed with speed in mind](#) (a.k.a. low-latency protocols, as indicated by 46% of respondents).

This is the only way to get streams into the sub-three-second range, but requires vendor support across the CDN and player.



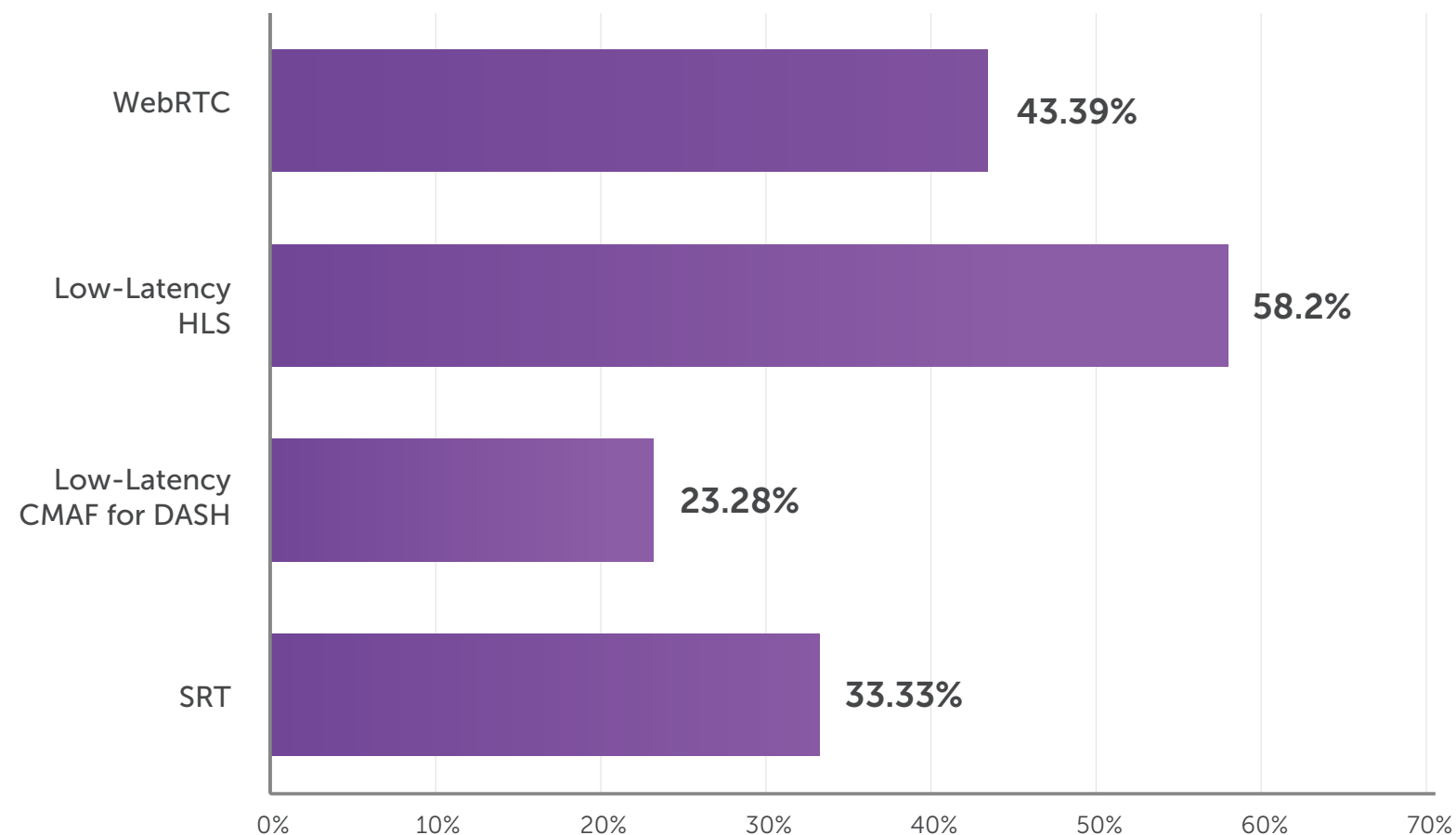
# Q12 Which low-latency technologies do you plan to use in the future?

## FUTURE TECHNOLOGIES



Over the past ten years, the streaming industry has been hard at work driving down latency. Several cutting-edge [protocols](#) — including [WebRTC](#), [Low-Latency HLS](#), [Low-Latency CMAF for DASH](#), and [SRT](#) — have come about from these efforts.

Here's a look at how they compare.



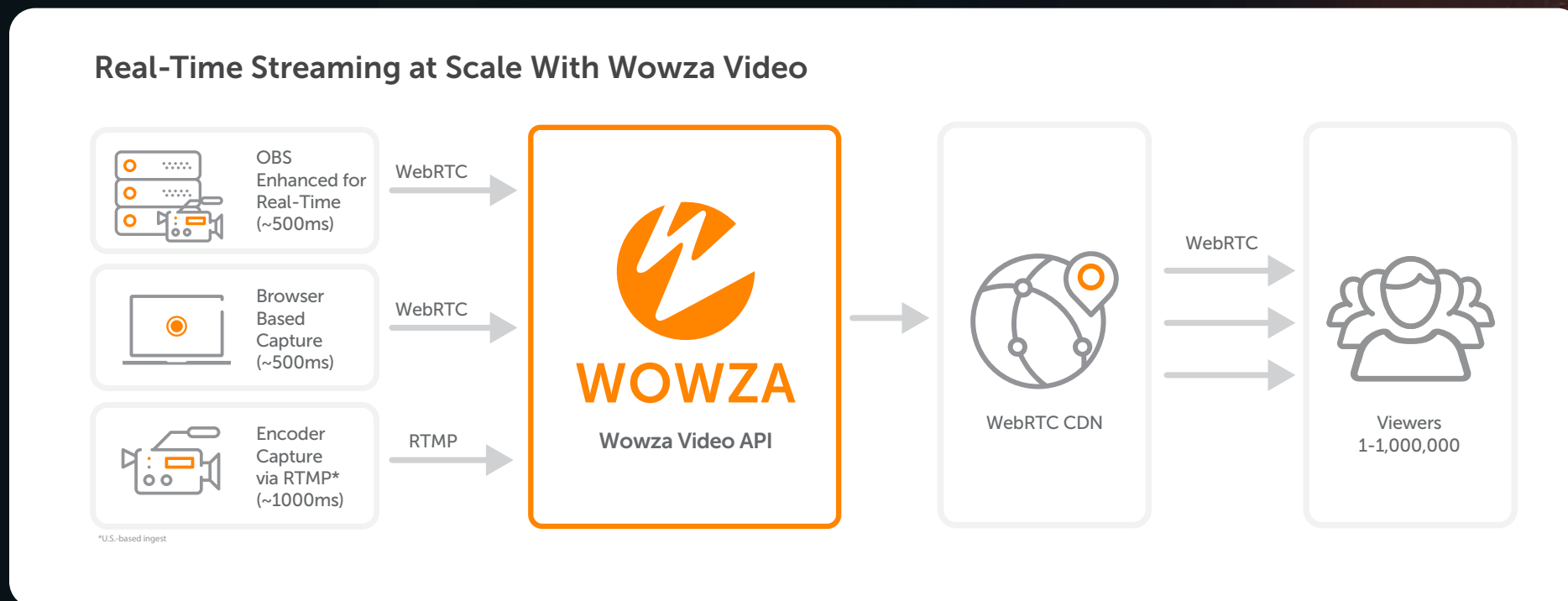
PROTOCOL	BENEFIT	LIMITATION
WebRTC	Real-time interactivity without a plugin.	Difficult to scale without a streaming platform like Wowza.
Apple Low-Latency HLS	Sub-three-second streaming supported by Apple.	Large-scale deployments aren't yet commonplace.
Low-Latency CMAF for DASH	A standard HTTP format that can be leveraged for low latency.	The introduction of Low-Latency HLS stifled momentum.
SRT	High-quality, low-latency video over suboptimal networks.	Primarily used for video contribution.



# Conclusion

For applications ranging from gambling and bidding to user-generated content, there's simply no room for latency. Today's broadcasters need speed and quality, without sacrificing audience size.

We designed our [Real-Time Streaming at Scale feature](#) for [Wowza Video](#) for sub-second delivery to viral audiences – ensuring a time-synchronized experience across the globe. Our platform instantly scales to accommodate even the biggest surprises. Plus, our team of Wowza Professional Services engineers stands ready to build a custom solution for your business needs.



Ready to learn more about Real-Time Streaming at Scale?

**[Contact us today!](#)**



### About Wowza

Wowza is the global leader in live streaming video. Our full-service platform powers reliable, secure, low-latency video delivery for companies worldwide. With more than a decade of experience working with 35,000+ organizations in industries ranging from media and entertainment to healthcare and surveillance, Wowza provides the performance and flexibility that today's businesses require.

We work with each customer to ensure their success in putting streaming to work for their business.

**If you can dream it, Wowza can stream it.**