VIVO CONTROL

TECHNOLOGY · 14 PAGES

WebRTC and the Modern Video Intercom — What Property Managers Need to Know

A plain-language explanation of how WebRTC-based video intercoms work, why they're more reliable than SIP systems, and what to look for in a vendor's technical infrastructure.

14 pages · 12 min read · PDF

INTRODUCTION

Why the Underlying Technology Matters

Every video intercom vendor will tell you their system "just works." Whether that's actually true depends on a piece of technology most property managers have never had a reason to ask about: how the video call itself gets from the panel at your gate to the resident's phone.

There are two fundamentally different ways to build that connection — an older approach called SIP, and a newer approach called WebRTC. You don't need to become an engineer to understand the difference, but understanding it changes what questions you ask in a vendor demo, and it explains why some video intercoms feel instant and reliable while others lag, drop calls, or require complicated on-site equipment.

> **WHAT YOU'LL BE ABLE TO DO AFTER READING THIS**
>
> Explain, in plain terms, why a WebRTC-based system tends to be more reliable than a SIP-based one — and ask any vendor the specific technical questions that reveal which one they've actually built.

THE BASICS

What Is WebRTC?

WebRTC stands for Web Real-Time Communication. It's an open standard, built and maintained jointly by the organizations that set the technical rules for the web, that lets two devices — say, a video intercom panel and a resident's smartphone — connect directly for live audio and video, without either side needing to install special software.

If you've ever made a video call directly inside a web browser with no app download, or joined a video meeting by clicking a link instead of installing anything, you've used WebRTC. It's the same technology, adapted for building entry.

Why It Was Built

Before WebRTC became widely supported (roughly 2013–2020, as browsers and mobile operating systems adopted it), real-time video required either a native app or specialized telephony equipment. WebRTC was designed specifically to make real-time video and audio a built-in capability of the browser and mobile operating system itself — which is exactly why it turned out to be a strong fit for an intercom that needs to reach a resident wherever they happen to be.

THE BASICS

What Is SIP, and How Do Older IP Intercoms Use It?

SIP stands for Session Initiation Protocol. It's an older standard, originally built for internet-based phone systems (VoIP), that a generation of IP-based intercoms adopted starting in the early 2000s. SIP intercoms typically place a call the way a traditional phone system would: through a dedicated call-routing server, often located on-site or managed by an IT provider.

Why SIP Intercoms Can Be More Complex to Run

A SIP-based system generally needs a properly configured network to work reliably — things like open ports, correctly configured routers, and sometimes a static IP address or an on-premises server. None of this is impossible to set up, but it typically requires more specialized IT involvement, both at installation and any time something changes on the network afterward.

To be clear: SIP is not "bad" technology, and plenty of IP intercoms built on it work fine in a well-maintained network environment. The relevant question for a property manager isn't whether SIP works — it's whether your property has the on-site IT support to keep a SIP-dependent system configured correctly over years of network changes, staff turnover, and equipment upgrades.

COMPARISON

WebRTC vs. SIP, Side by Side

Factor	SIP-Based Intercom	WebRTC-Based Intercom
Network setup	Often requires port forwarding, static IP, or on-premises server	Works over standard internet or mobile data, no special configuration
Works on cellular networks	Frequently unreliable behind carrier networks	Designed to work over 4G/5G mobile data reliably
Resident app requirement	Often a dedicated, proprietary app	Works natively in browsers and standard apps
On-site IT dependency	Higher — network misconfiguration is a common failure point	Lower — cloud backend handles connection logic
Encryption	Varies by implementation	Encrypted by default (DTLS-SRTP) as part of the standard
Typical latency	Can vary, sometimes noticeably delayed	Sub-second, real-time by design

Not every SIP intercom performs at the low end of this table, and not every WebRTC platform is implemented well — the underlying protocol is necessary but not sufficient. Use the vendor questions in this guide to verify actual implementation quality, not just the buzzword.

RELIABILITY

Why WebRTC Tends to Be More Reliable in Practice

Three specific technical characteristics of WebRTC translate directly into a better day-to-day experience for residents and staff.

It Doesn't Depend on a Static Network Setup

A resident might be on their building's WiFi one moment and on cellular data the next, walking from the lobby to the parking lot. WebRTC is built to establish and maintain a connection across exactly this kind of change, without a dedicated on-premises server keeping track of it.

It Was Designed for Mobile Networks from the Start

SIP predates the smartphone era and was originally built with wired office phone networks in mind. WebRTC was standardized during the smartphone era specifically to handle mobile data conditions — variable signal strength, network switching, and carrier-level restrictions that often cause SIP-based calls to drop or fail to connect at all.

It Removes the On-Premises Point of Failure

A SIP system that depends on an on-site server or specific router configuration has a single point of failure sitting in your building — if that hardware fails, is misconfigured after a network change, or isn't maintained, the intercom stops working. A cloud-native WebRTC platform moves that responsibility off your property entirely.

SECURITY

Encryption: What's Built In vs. What's Optional

One of the most consequential technical differences between WebRTC and SIP has nothing to do with speed — it's encryption.

WebRTC's technical specification requires encryption for every media stream by default, using a combination called DTLS-SRTP. This isn't a feature a vendor chooses to add — it's part of what makes something WebRTC in the first place. Every video and audio stream between a WebRTC-based intercom panel and a resident's phone is encrypted as a baseline requirement of the technology.

SIP, by contrast, does not encrypt by default. Encryption on a SIP system is possible, but it depends on the specific implementation, and it's the kind of thing that can be configured correctly, configured incorrectly, or simply skipped under time pressure during installation.

> **THE QUESTION TO ASK**
>
> Don't ask "is it encrypted?" — ask "is encryption built into the protocol, or is it something your team configures separately?" The answer tells you whether security is structural or optional.

RESIDENT EXPERIENCE

What This Actually Looks Like for a Resident

All of this technical difference shows up in a handful of very concrete, everyday moments.

- **A visitor scans a QR code and the video call connects immediately** — no "connecting..." delay, no dropped first attempt.
- **A resident answers the intercom call while driving home, switching from WiFi to cellular data mid-call** without the call cutting out.
- **A resident answers directly in their phone's browser or a lightweight app** — no separate proprietary client to install, update, or troubleshoot.
- **The call quality stays consistent** even on an average residential internet connection or a moderate cellular signal, because the underlying protocol was built for exactly those variable conditions.
- **Nothing breaks when the property's internet router is replaced or reconfigured,** because there's no on-premises SIP server whose settings need to be re-applied.

EVALUATION

What to Look for in a Vendor's Technical Infrastructure

- **Cloud-native backend.** The system's call-routing and management logic should live in the cloud, not on a server physically located at your property.
- **No dedicated SIP server requirement.** If a vendor tells you that you'll need on-site networking equipment specifically to route calls, that's a SIP-style dependency worth asking more about.
- **Native browser and mobile app support.** Residents shouldn't need to install and maintain a proprietary calling app just to answer the intercom.
- **Encryption as a stated technical standard,** not a feature added on request — ask specifically whether it's DTLS-SRTP or an equivalent standard applied to every call by default.
- **Confirmed performance on mobile data,** not just building WiFi — ask for a live demo specifically over cellular data, not just on the vendor's office network.
- **Field-updatable hardware.** A panel that receives firmware updates remotely, without a technician visit, points to a platform built to be maintained centrally rather than one-off per property.
- **A fallback path.** Ask what happens if internet service to the property goes down entirely — a well-built platform typically has a documented fallback, such as routing to a phone line (PSTN) as a backup path.

VENDOR QUESTIONS

Questions to Ask About Technical Infrastructure

1. Is your video calling built on WebRTC, SIP, or a proprietary protocol? If SIP, what on-site networking does it require?
2. Do residents need to install a dedicated app, or does it work in a standard mobile browser or lightweight app?
3. Is encryption built into the protocol by default, or configured separately? What specific standard is used?
4. Can you demonstrate a call live, over cellular data, walking away from the building's WiFi mid-call?
5. What happens to call quality and reliability if the property's internet service goes down or is replaced?
6. Is the panel's firmware updated remotely, or does it require an on-site technician visit?
7. Is there a documented fallback path (such as a phone-line connection) if internet connectivity fails entirely?

A vendor confident in their infrastructure will answer these specifically and immediately. Vague or evasive answers to any of these are worth treating as a flag, not a formality.

CONTEXT

A Brief History: How Intercom Technology Got Here

WebRTC-based video intercoms are the latest step in a much longer evolution — understanding that arc makes it clear why this shift happened, rather than treating it as an arbitrary buzzword.

Era	Technology
1910s–1950s	Mechanical speaker tubes and electric buzzers
1950s–1980s	Hardwired two-wire audio intercoms
1980s–2000s	Analog, then color, video intercoms with in-unit monitors
2000s–2012	IP-based intercoms over ethernet, using the SIP protocol
2012–2019	Smartphone-connected intercoms, calling a resident's mobile app
2019–2023	Cloud-managed access platforms with remote administration
2023–present	WebRTC-native platforms with cloud backends and no on-site server requirement

Each step solved a real limitation of the one before it. WebRTC's role in this timeline is specifically to remove the on-premises networking complexity that SIP-based IP intercoms introduced — while keeping everything those systems got right about video verification and remote answering.

CLEARING THINGS UP

Common Misconceptions

- **"It's an IP intercom, so it must be modern."** IP just means the system communicates over a network instead of dedicated copper wiring — it says nothing about whether it uses SIP or WebRTC underneath. Many "IP intercoms" on the market today are SIP-based systems from the 2000s IP era, not WebRTC platforms.
- **"WebRTC is a brand, not a standard."** WebRTC isn't owned or branded by any single vendor — it's an open standard maintained by the same bodies that govern core web technology. Any vendor can build on it, and asking whether they do is a fair, neutral technical question.
- **"SIP systems never work well."** That's not accurate either — a properly maintained SIP system with dedicated IT support can perform reliably. The real issue is that this maintenance burden falls on your property, ongoing, rather than being handled centrally by the vendor's cloud infrastructure.
- **"This is only relevant to the IT department."** Call reliability directly affects whether a resident actually answers the intercom for a delivery or guest — which is a resident satisfaction and security issue, not just a technical one.

CONCLUSION

Key Takeaways

- WebRTC and SIP are two different ways of building the connection behind a video intercom call — the choice affects reliability, security, and IT complexity, not just call quality.
- WebRTC was built for the mobile, browser-native era and generally performs better across changing networks and cellular data without on-premises equipment.
- Encryption is a default requirement of WebRTC, while it's optional and implementation-dependent on SIP.
- "IP intercom" is not the same claim as "WebRTC intercom" — ask specifically which protocol underlies any system you're evaluating.
- Use the seven vendor questions in this guide during any demo, and ask to see a live call over cellular data, not just building WiFi.

Pair this guide with our Access Control RFP Template — Section 4's technical requirements and Section 6's technical proposal questions are a natural place to formalize the questions covered here.

VIVO CONTROL

Built WebRTC-Native, From the Ground Up

Vivo Control's video intercom runs on a WebRTC-native, cloud-managed architecture — encrypted by default, no on-premises server required, and reliable performance across WiFi and cellular data alike.

Book a Free Demo →

---

www.vivocontrol.com/en/demo/ · hello@vivocontrol.com · This guide is for general educational purposes and simplifies technical concepts for a non-engineering audience.