

WHITEPAPER

# Verzögerung in der netzwerkbasieren Live-Videosicherheit

Juni 2024

# Zusammenfassung

Live-Streaming bei der IP-Videosicherheit besteht aus dem Aufzeichnen des Videos mit der Kamera, dem Packen und dem Transport der Videodaten durch das Netzwerk sowie dem Entpacken beim Empfänger zur Anzeige auf dem Bildschirm. Jeder dieser Schritte verursacht mehr oder weniger Verzögerung.

- **Von der Kamera verursachte Verzögerung.** Auf das Aufzeichnen von Videomaterial folgt die Bildverarbeitung, Komprimierung und das Packen. Jeder Schritt führt zu einer gewissen Verzögerung, aber insgesamt machen die Vorgänge innerhalb der Kamera nur einen kleinen Teil der gesamten End-to-End-Verzögerung aus.
- **Vom Netzwerk verursachte Verzögerung.** Sie kann sowohl sehr hoch als auch sehr gering ausfallen und ist der unberechenbarste Faktor in der „Gleichung“ für die Verzögerung. Berechenbarer wird sie, indem Sie in gute Netzwerksoftware und -hardware investieren. Die Netzwerklatenz hängt sehr stark vom Verhältnis zwischen Daten und Bandbreite ab. Sie können die Kamera so konfigurieren, dass die von ihr erzeugte Datenmenge reduziert wird, womit sich auch die Anzahl der über das Netzwerk zu sendenden Pakete verringert.
- **Von der Client-Seite verursachte Verzögerung.** Auf der Client-Seite werden die Daten empfangen und gepuffert, um dann sortiert und in eine Warteschlange zur Grafikkarte und zum Monitor geleitet zu werden. Die Verzögerung wird vor allem durch den Empfangspuffer im Client beeinflusst, dessen Latenz mehrere Sekunden betragen kann. Ohne einen großen Puffer besteht die Gefahr, dass der Videostream nicht gleichmäßig wiedergegeben wird.

Die Reduzierung der Verzögerung ist immer eine Frage der Kosten. Die größten Optimierungen lassen sich durch die Verbesserung des Netzwerks und der clientseitigen Hard- und Software erzielen.

# Inhalt

1	Einführung	4
2	Was ist Verzögerung?	4
3	Wie wird Verzögerung gemessen?	4
4	Welche Faktoren beeinflussen die Verzögerung?	5
4.1	Von der Kamera verursachte Verzögerung	5
4.2	Netzwerklatenz	7
4.3	Von der Client-Seite verursachte Verzögerung	10
5	Reduzierung der Verzögerung	11
5.1	Reduzierung der Verzögerung in der Kamera	11
5.2	Reduzierung der Verzögerung im Netzwerk	12
5.3	Reduzierung der Verzögerung auf der Client-Seite	13

# 1 Einführung

Im Kontext der Videosicherheit ist die Verzögerung die Zeitspanne zwischen dem Zeitpunkt der Aufnahme eines Bildes und dem Zeitpunkt der Anzeige desselben Bildes. Sie wird auch als End-to-End-Verzögerung oder Sensor-zu-Bildschirm-Verzögerung bezeichnet. Die Übertragung eines Bildes vom Kamerasensor zum Bildschirm besteht aus einer langen Kette von Schritten.

In diesem Whitepaper werden die verschiedenen Schritte beschrieben, aus denen sich die Gesamtverzögerung ergibt. Außerdem werden Empfehlungen zur Reduzierung der Verzögerung gegeben.

## 2 Was ist Verzögerung?

Die Definition von Verzögerung hängt vom jeweiligen Kontext ab. In der Netzwerk-Technologie wird die Verzögerung oder auch Latenzzeit gemeinhin als die Zeitspanne zwischen dem Senden einer Information von der Quelle und dem Empfang derselben Information am endgültigen Zielort verstanden.

In diesem Whitepaper geht es um die Verzögerung in Netzwerk-Videosicherheitssystemen. Die Verzögerung ist hier die Zeitspanne zwischen der Aufnahme eines Bildes durch eine Kamera und dem Zeitpunkt der Anzeige auf einem Bildschirm. In dieser Zeit wird das Bild aufgenommen, komprimiert, übertragen, dekomprimiert und angezeigt. Jeder Schritt trägt seinen eigenen Anteil zur Gesamtverzögerung bei. Vereinfacht kann bei der Ermittlung der End-to-End-Verzögerung von drei Hauptphasen ausgegangen werden:

- Von der Kamera verursachte Verzögerung (Bildverarbeitungs- und Komprimierungslatenz)
- Vom Netzwerk verursachte Verzögerung (Übertragungslatenz)
- Von der Empfängerseite verursachte Verzögerung (Puffer-, Dekomprimierungs- und Anzeigelatenz auf der Client-Seite)

Um die anvisierte maximale Verzögerung eines Videosicherheitssystems zu erreichen, muss jede dieser Latenzen bei der Projektierung der Videolösung berücksichtigt werden.

## 3 Wie wird Verzögerung gemessen?

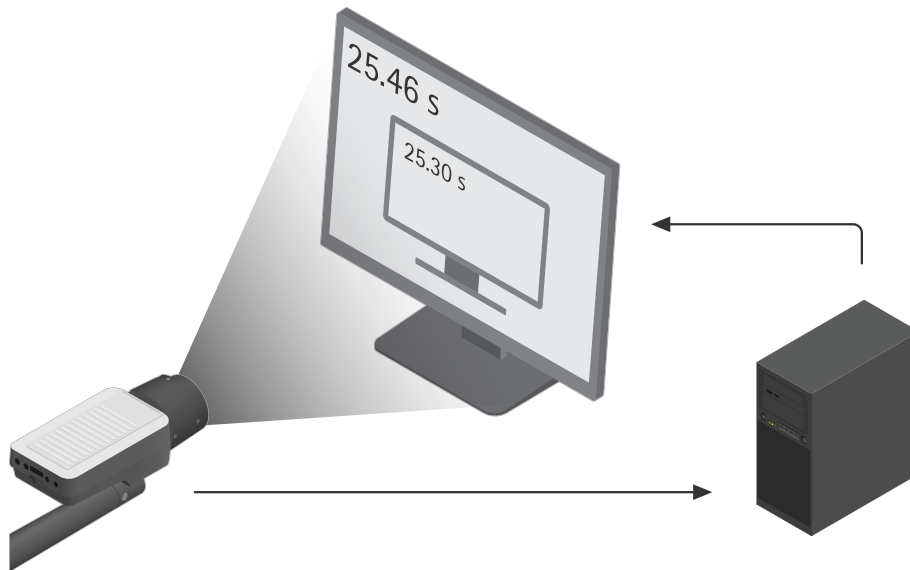
Die Verzögerung wird üblicherweise in Zeiteinheiten ausgedrückt, in der Regel in Sekunden (s) oder Millisekunden (ms). Eine genaue Messung der Verzögerung ist äußerst schwierig, da dazu die Uhren in Kamera und Bildschirm exakt synchronisiert werden müssten. Eine einfache Methode (mit dem Vorbehalt einer minimalen Abweichung von den exakten Werten) ist die Nutzung der Overlay-Text-Funktion für Zeitstempel. Diese Methode misst die Verzögerung eines Videosicherheitssystems, d. h. die Zeitspanne zwischen der Aufnahme eines Bildes mit dem Objektiv und der Darstellung desselben Bildes auf dem Bildschirm eines Überwachungsgeräts.

Dabei ist zu beachten, dass bei dieser Methode ein Fehler von bis zu einem Bildintervall auftreten kann. Grund hierfür ist, dass die zur Berechnung der Verzögerung verwendeten Zeitstempel erst bei Aufnahme eines Bildes erfasst werden. Daher lässt sich die Verzögerung also nur mit dem Faktor der Bildrate berechnen. Bei einer Bildrate von 25 Bildern pro Sekunde kann die Verzögerung demnach als ein Vielfaches von 40 ms berechnet werden. Beträgt die Bildrate 1 Bild pro Sekunde, wird die Verzögerung als ein Vielfaches von Sekunden berechnet. Diese Methode wird daher nicht für niedrige Bildraten empfohlen.

Messung der Verzögerung mit der Overlay-Text-Funktion für Zeitstempel:

1. Schalten Sie mit Hilfe von (%T:%f) den Zeitstempel im Overlay ein.

2. Platzieren Sie die Kamera so, dass sie die Ausgabe ihres eigenen Livestreams erfasst.
3. Erstellen Sie Schnappschüsse von der Livestream-Ausgabe und vergleichen Sie die Zeitangaben im ursprünglichen Text-Overlay und in der Bildschirmschleife.



*Ein Beispiel für die Messung der Verzögerung mithilfe der Overlay-Text-Funktion für Zeitstempel. Wir können einen Zeitunterschied von 25,46 bis 25,30 = 0,16 Sekunden bzw. 160 ms feststellen. Die End-to-End-Verzögerung beträgt also 160 ms.*

## 4 Welche Faktoren beeinflussen die Verzögerung?

Die gesamte Verzögerung ist die Summe der Verzögerungen, die durch Kamera, Netzwerk und Client-Seite verursacht werden.

### 4.1 Von der Kamera verursachte Verzögerung

Für jedes Bild fällt eine Belichtungszeit von etwa  $1/30$  s an, gefolgt von einer kurzen Zeit zum Skalieren und Kodieren des Bildes. Das kodierte Bild wird geteilt und gepackt, wonach alle 33 ms ein Bild in das Netzwerk ausgegeben wird. Die Zeit, die für diesen Vorgang in der Kamera benötigt wird, kann weniger als 50 ms betragen, liegt aber in der Regel bei einigen hundert ms. Sie variiert leicht je nach Kamera (PTZ ausgeschlossen) und Art des Bildes (I-Frame oder P-Frame).

#### 4.1.1 Aufnahmeverzögerung

Werfen wir einen Blick ins Innere der Videokamera. Der Bildsensor der Kamera besteht aus Millionen von Photodetektoren (lichtempfindlichen Punkten), den sogenannten Pixeln. Der Sensor fängt während eines Belichtungsintervalls Licht in seinen Pixeln ein, bevor er die registrierte Lichtenergie in elektronische Signale umwandelt. Nach der Umwandlung sind die Pixel wieder frei und bereit für eine weitere Belichtung. Die Anzahl der Belichtungen des Sensors pro Zeiteinheit, also die Anzahl der Bilder, die die Kamera pro Sekunde aufnehmen kann, richtet sich nach der Aufnahmezeit des Sensors.

Die Aufnahmeverzögerung hängt von der Aufnahme­rate ab. Wenn Sie die Aufnahme­rate auf 30 Bilder pro Sekunde einstellen, was bedeutet, dass der Sensor alle 1/30 Sekunde ein Bild aufnimmt, beträgt die Verzögerung bis zu 33,3 ms.

#### 4.1.2 Durch die Bildverarbeitung verursachte Verzögerung

Jedes erfasste Bild durchläuft eine Reihe von Bildverarbeitungsschritten, wie z. B. Deinterlacing, Skalierung und Bilddrehung, die eine zusätzliche Verzögerung verursachen. Je mehr das Video durch die Kamera verarbeitet wird, desto größer die Verzögerung. Da aber die Verarbeitung in der Kamera die Menge der erzeugten Daten beeinflusst, wirkt sich der Umfang der Verarbeitung im Weiteren auch auf die Verzögerung bei der Übertragung der Daten über das Netzwerk aus.

Einige Parameter, die die Verzögerung beeinflussen, sind:

- **Bilddrehung.** Die Drehung des Videostreams um 90 oder 270 Grad verursacht zusätzliche Last für den Kodierungsprozessor. Pixel müssen umgeordnet und gepuffert werden, was zu einer gewissen Verzögerung führt.
- **Auflösung.** Je höher die Auflösung, desto mehr Pixel muss der Prozessor kodieren. Der Anstieg der Verarbeitungszeit bei einer höheren Auflösung ist im Vergleich zu einer niedrigeren Auflösung in der Regel unbedeutend, da er in hochauflösenden Kameras durch einen schnelleren Prozessor ausgeglichen wird. Eine höhere Auflösung führt jedoch zu mehr Daten pro Bild und damit zu einer höheren Anzahl zu übertragender Pakete. In einem Netzwerk mit begrenzter Bandbreite kann dies Verzögerungen bei der Übertragung nach sich ziehen. Das führt wiederum dazu, dass auf der Empfängerseite ein größerer Puffer benötigt wird, wodurch die Verzögerung steigt.
- **Rauschfilterung.** Die hochentwickelte Rauschfilterung erfordert das Puffern mehrerer Bilder, was zu einer zusätzlichen Verzögerung führt.
- **Privatzonenmaskierung.** Erweiterte Funktionen zur Maskierung von Privatzonen, wie AXIS Live Privacy Shield, können zusätzliche Verzögerungen verursachen. Das liegt an der Pufferung, die erforderlich ist, um sicherzustellen, dass die richtigen Privatzonenmasken zum richtigen Zeitpunkt angewendet werden.

#### 4.1.3 Komprimierungs­latenz

Videos werden kodiert, um die Daten vor der Übertragung zu komprimieren. Bei der Komprimierung werden nach einem oder mehreren mathematischen Algorithmen Bilddaten entfernt. Je nach der Menge der zu verarbeitenden Daten nimmt dies mehr oder weniger Zeit in Anspruch. Die Komprimierungs­latenz in diesem Schritt wird durch verschiedene Aspekte der Komprimierung beeinflusst:

- **Komplexität der Komprimierungsalgorithmen**

H.264 und H.265 sind höherentwickelte Komprimierungsverfahren als MJPEG. Axis Kameras verfügen im Vergleich zur MJPEG-Komprimierung jedoch in der Regel über eine höhere Kapazität für die H.264- und H.265-Komprimierung, was bedeutet, dass die Komprimierungs­latenz bei H.264 oder H.265 nicht unbedingt höher ist. Auf der Dekodierungsseite kann sie allerdings höher ausfallen, weil der Decoder für die H.264- und H.265-Datenströme, die von Axis Kameras erzeugt werden, mindestens ein Bild puffern muss, während die MJPEG-Dekodierung keinen Puffer benötigt. Darüber hinaus fällt beim Zipstream-Speicherprofil eine zusätzliche Verzögerung von bis zu zwei Bildern pro Sekunde an, d. h. 66,7 ms für ein Video mit 30 Bildern pro Sekunde.

- **Effektivität des Komprimierungsverfahrens**

Die in Axis Kameras am häufigsten verwendeten Kodierungsverfahren sind MJPEG, H.264 und H.265. Sie alle führen zu einer Verzögerung in der Kamera. H.264 und H.265 minimieren den Videodurchsatz

stärker als MJPEG, was bedeutet, dass weniger Datenpakete über das Netzwerk gesendet, entpackt und beim Empfänger gerendert werden müssen. Dadurch wird die Gesamtverzögerung reduziert.

- **Wahl der Bitrate**

Die Videokomprimierung verringert die Größe der Videodaten. Allerdings sind die Bilder nach der Komprimierung unterschiedlich groß. Die Größe der komprimierten Daten kann je nach Szene variieren. Mit anderen Worten bestehen die komprimierten Originaldaten aus Videostreams mit variabler Bitrate (VBR), so dass eine variable Bitrate in das Netzwerk ausgegeben wird. Hierbei sind die Beschränkungen des verfügbaren Netzwerks zu berücksichtigen, z. B. Bandbreitenbeschränkungen. Eine begrenzte Bandbreite des Videostreaming-Systems erfordert in der Regel eine Regulierung der Übertragungsbitrate. Bei einigen Encodern können Sie zwischen VBR und der maximalen Bitrate (MBR) wählen. Wenn Sie sich für MBR entscheiden, ist gewährleistet, dass das Netzwerk nur eine begrenzte Datenmenge erhält. Indem Sie eine Überlastung des Netzwerks vermeiden, verringern Sie die Netzwerkverzögerung und den Bedarf an einem größeren Puffer auf der Empfängerseite des Systems.

Bei Axis Kameras bieten H.264- und H.265-Encoder die Möglichkeit, zwischen VBR und MBR zu wählen. Wir empfehlen jedoch generell die Verwendung von VBR bei Netzwerk-Videos, bei denen die Qualität in Echtzeit an den Szeneninhalte angepasst wird. Es wird davon abgeraten, MBR grundsätzlich zur Speicherreduzierung oder zum Ausgleich schwacher Netzwerkverbindungen zu nutzen, da Kameras, die MBR-Videos liefern, in kritischen Situationen gezwungen sein könnten, wichtige forensische Details zu löschen.

Die Verwendung von Zipstream in der Kamera reduziert die Bitrate. Hierdurch wird die Datenmenge begrenzt und die Verzögerung im Netzwerk verringert.

- **Anzahl der Videostreams.** Wenn verschiedene Videostreams von der Kamera angefordert werden (unterschiedliche Bildraten oder Auflösungen), führt die Verarbeitung der zusätzlichen Videostreams zu einer weiteren Verzögerung, da alle Videostreams von demselben Prozessor kodiert werden müssen.

Bei der Wahl des Komprimierungsverfahrens sollten Sie alle diese Aspekte berücksichtigen. Einerseits kann ein hochentwickelter Encoder-Algorithmus eine längere Zeit für die Kodierung und Dekodierung in Anspruch nehmen, andererseits wird dadurch die über das Internet gesendete Datenmenge reduziert, was wiederum die Übergangszeiten verkürzt und die Größe des Puffers auf der Empfängerseite verringert.

#### **4.1.4 Pufferlatenz**

Da Videos Bild für Bild verarbeitet werden, kann nur eine begrenzte Menge an Daten in einem Vorgang komprimiert werden. Zwischen den Verarbeitungsschritten sind manchmal Kurzzeitpuffer erforderlich, die eine zusätzliche Verzögerung in der Kamera bedeuten.

#### **4.1.5 Audiolatenz**

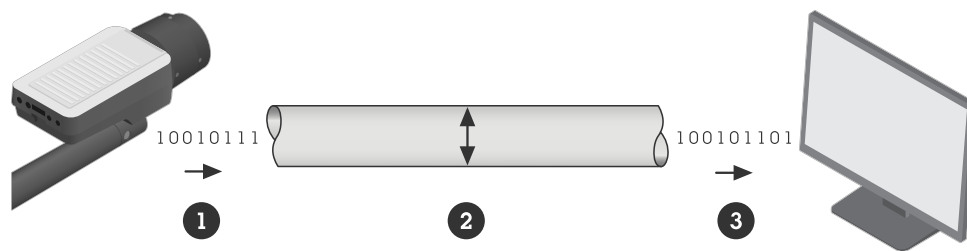
In einigen Fällen wird der Videostream von Audio begleitet. Der Audio-Encoder muss eine bestimmte Anzahl von Samples abwarten, bevor ein Block zum Beginn der Audiocodierung verfügbar ist, was zu einer zusätzlichen Verzögerung auf der Kameraseite führt. Die verschiedenen Audio-Kodieralgorithmen haben unterschiedliche Abstraten und Blockgrößen.

## **4.2 Netzwerlatenz**

Nachdem das Bild erfasst, verarbeitet und komprimiert wurde, werden die Videodaten über ein Netzwerk übertragen, bevor sie auf der Client-Seite empfangen und gerendert werden. Um zu verstehen, wie sich das Netzwerk auf die Verzögerung auswirkt, müssen wir zunächst folgende grundlegende Konzepte der

Video-Netzwerktechnologie betrachten: Bitrate, Bandbreite und Durchsatz. Die Netzwerklatenz verhält sich proportional zur Bitrate und umgekehrt proportional zur Bandbreite.

- Die **Bitrate** ist die Datenmenge in Bits, die pro Zeiteinheit verarbeitet wird. Bei der Videosicherheit ist die Bitrate definiert als die Menge der von der Kamera erzeugten Daten, die pro Gerät über das Netzwerk gesendet werden. Die Bitrate hängt von vielen Faktoren ab, sehr stark z. B. von der aufgenommenen Szene, der Verarbeitung in der Kamera und den Einstellungen für den Videostream. Wenn die Kamera mehr zu übertragende Daten erzeugt, ist bei einer begrenzten Bandbreite mit einer höheren Netzwerklatenz zu rechnen.
- Die **Bandbreite** ist die Menge der Daten, die das Netzwerk potenziell zwischen Kamera und Monitor übertragen kann. Es handelt sich also um die maximale Übertragungskapazität Ihrer Verbindung. Die Bandbreite hängt von der Länge und der Infrastruktur der Verbindung ab, d. h. von Switches, Routern, Kabeln, Proxys usw. Wenn wir die Kapazität des Netzwerks erhöhen, können mehr Daten übertragen werden, was zu einer geringeren Verzögerung führt.
- Der **Durchsatz** ist die tatsächlich erreichte Geschwindigkeit Ihrer Datenübertragung (in Bits/s). Er ist davon abhängig, ob Sie die Verbindung mit anderen teilen und wie stark die elektromagnetischen Störungen in den Verbindungskabeln sind. Außerdem kann der Durchsatz durch die für die Ports gewählten QoS-Einstellungen (Quality of Service – Servicequalität) begrenzt werden.



Wenn wir die Verbindung (das Netzwerk) zwischen der Kamera und dem Monitor mit einem Rohr vergleichen würden, wäre die Bitrate (1) ein Maß dafür, wie viele Daten pro Zeiteinheit in das Rohr fließen, die Bandbreite (2) wäre der Durchmesser des Rohrs und der Durchsatz (3) entspräche der Datenmenge pro Zeiteinheit, die tatsächlich durch das Rohr fließt.

- 1 Bitrate
- 2 Bandbreite
- 3 Durchsatz

Die Gesamtnetzwerklatenz hängt von drei wichtigen Faktoren ab: der Infrastruktur der Verbindung zwischen der Kamera und dem Videoanzeigerät (was die Bandbreite bestimmt), der von der Kamera erzeugten Datenmenge (was die Bitrate bestimmt) und der Wahl des Übertragungsprotokolls.

#### 4.2.1 Die Infrastruktur

Das Netzwerk ist der Faktor für die End-to-End-Verzögerung, der sich am wenigsten vorhersehen lässt. Switches, Router, Kabel, Proxys – alles, was sich im Netzwerk zwischen Sender und Empfänger befindet, wirkt sich auf die Gesamtverzögerung aus. In einem LAN kann die Netzwerklatenz nur wenige ms betragen, womit sie unbedeutend niedrig und vernachlässigbar ist. Wenn der Videostream jedoch über das Internet auf nicht spezifizierten Routen übertragen werden soll, ist die Netzwerklatenz schwer vorhersehbar und könnte in vielen Fällen der Hauptfaktor für die End-to-End-Verzögerung sein.

Mit einer sorgfältigen Verwaltung des Netzwerks und der Zuweisung der Bandbreite lassen sich die unvorhersehbaren Quellen für Netzwerklatenz vorhersagbarer machen. Die Verbindung zwischen Kamera und Anzeigerät muss einen garantierten Durchsatz haben.



In einem LAN (Local Area Network) kann dies realisiert werden, indem sichergestellt wird, dass die Verbindung möglichst wenige Netzwerk-Sprünge (Hops) enthält. Die Verbindung sollte nicht mit anderem Datenverkehr wie Voice over IP (VoIP) oder anderen Protokollen geteilt werden, die standardmäßig Vorrang vor Video haben, oder mit anderen anspruchsvollen Diensten, die die Verbindung überlasten.

Wenn die Verbindung über ein WAN (Wide Area Network) erfolgt, muss bei jedem Netzwerk-Sprung (also in den Routern und Switches) die QoS gewährleistet sein. Hierfür können Sie über Ihren lokalen Internetanbieter eine Punkt-zu-Punkt-Verbindung mieten.

Konfigurierbare Faktoren, die den Durchsatz beeinflussen:

- Paket-Overhead (protokollabhängiges Beispiel: VLAN-Header)
- Proxys und Firewalls
- QoS der einzelnen Verbindungen auf der gesamten Netzwerkroute
- Burst-Modus oder nicht (aktiviert → höhere Geschwindigkeit)
- MTU – die Größe der Video-Nutzdaten

Kostenbezogene Faktoren, die den Durchsatz beeinflussen:

- Prozessorgeschwindigkeit und Port-Puffer der Switches und Router
- Kabelart (oder Funkstandard)

#### 4.2.2 Die Menge gestreamter Videodaten

Das gewählte Bildverarbeitungs- und -komprimierungsverfahren in der Kamera wirkt sich auf die Netzwerklatenz aus, da diese Optionen die Menge der erzeugten Videodaten beeinflussen. Die Übermittlung einer geringeren Datenmenge erfordert logischerweise weniger Zeit.

#### 4.2.3 Die Übertragungsprotokolle

Die Videobilder der Kamera werden an ein Transportprotokoll, in der Regel RTP oder HTTP, übergeben. Die Übertragung zum Rendering-Client erfolgt über ein IP-Netzwerk. Zur Übertragung wird entweder das zuverlässige TCP, ein verbindungsorientiertes Protokoll mit Wiederübermittlung verlorener Datenpakete, oder UDP genutzt, ein einfacheres Protokoll, das keine Empfangsgarantie und keine Möglichkeit zur Wiederübermittlung verlorener Pakete bietet.

Axis Kameras bieten Ihnen die folgenden Optionen, um den kodierten Datenstrom für die Übertragung zu kapseln:

*Tabelle 4.1 Optionen für die Kapselung des kodierten Datenstroms.*

Topologie	Empfohlene Kapselungsmodi für Axis Videodatenpakete
LAN / weniger Netzwerk-Sprünge und direkt verwaltete Knotenpunkte	MJPEG / HTTP / TCP
LAN / weniger Netzwerk-Sprünge und direkt verwaltete Knotenpunkte	H.264, H.265 oder MJPEG / RTP / RTSP / HTTP / TCP
LAN / weniger Netzwerk-Sprünge und direkt verwaltete Knotenpunkte	H.264, H.265 oder MJPEG / RTP / RTSP / TCP

Tabelle 4.1. Optionen für die Kapselung des kodierten Datenstroms. (Fortsetzung)

Topologie	Empfohlene Kapselungsmodi für Axis Videodatenpakete
WAN / mehrere Netzwerk-Sprünge ohne volle Kontrolle über die Knotenpunkte	H.264, H.265 oder MJPEG / RTP / Unicast / UDP
WAN / mehrere Netzwerk-Sprünge ohne volle Kontrolle über die Knotenpunkte	H.264, H.265 oder MJPEG / RTP / Multicast / UDP
Remote-Verbindung / Cloud / WAN / mehrere Netzwerk-Sprünge ohne volle Kontrolle über die Knotenpunkte	H.264, H.265 oder MJPEG / RTP / WebRTC / UDP oder TCP

Normalerweise dauert der Transport eines Pakets über TCP länger als über UDP, da ein zusätzlicher Verbindungsaufbau, Bestätigungsnachrichten und die erneute Übertragung von Paketen bei Erkennen eines Verlusts erforderlich sind. Dagegen kann es bei UDP zu Artefakten oder Unterbrechungen im Videostream kommen, wenn Pakete während der Übertragung verloren gehen. TCP ist also dadurch geprägt, dass die Netzwerklatenz bei Paketverlusten über das übliche Maß steigen kann (Jitter), während bei UDP Artefakte und/oder Unterbrechungen bei Paketverlusten denkbar wären. Wenn Datenverluste und vorübergehende Qualitätseinbußen vertretbar sind, ist UDP eine mögliche Wahl für Netzwerke mit geringer Bandbreite.

Bei Nutzung von TCP müssen mehr Pakete gesendet werden, wozu Sie wiederum eine größere Bandbreite benötigen. Wenn Sie wissen, dass das Netzwerk stark ausgelastet ist, sollten Sie UDP als Übertragungsprotokoll wählen. Da Paketverluste hingenommen werden, können diese in der Folge auch auftreten und die Bildqualität beeinträchtigen.

Mit WebRTC und adaptiver Bitrate wird das Video an das Netzwerk angepasst, um unkontrollierbare Latenzspitzen zu vermeiden, wie sie bei TCP möglich sind.

### 4.3 Von der Client-Seite verursachte Verzögerung

Nachdem das Video auf der Client-Seite des Videosystems empfangen wurde, wird es entpackt, sortiert und dekodiert und anschließend mit einem Media-Player gerendert. Jeder Schritt trägt zur Gesamtverzögerung auf der Client-Seite bei, wo auch der Computer selbst eine wichtige Rolle spielt. Rechenkapazität, Betriebssystem, Netzwerkkarte und Grafikkarte wirken sich auf die Verzögerung aus. In der Regel ist MJPEG der Standard mit der geringsten Dekodier- und Anzeigelatenz, da die Daten ohne Zeitcodes auf dem Bildschirm angezeigt werden können. H.264 und andere Standards für die Videokomprimierung weisen jedem Bild Zeitcodes zu und verlangen, dass sie entsprechend gerendert werden.

#### 4.3.1 Ausspielpuffer

Reale Netzwerke sind oft sehr groß und kompliziert mit schwankender Datenauslastung und Paketen, deren Reihenfolge verändert wurde. Um die durch den Datentransport im Netzwerk verursachten Schwankungen auszugleichen, kommt auf der Client-Seite ein Puffer zum Einsatz. Dieser Puffer, der auch als Ausspielpuffer oder Jitter-Puffer bezeichnet wird, sorgt dafür, dass die Pakete in die richtige Reihenfolge gebracht werden, und puffert genügend Daten, damit der Decoder nicht leer läuft und die Bildrate im Viewer gleich bleibt.

Dieser Puffer verursacht eine vergleichsweise hohe Verzögerung. Verschiedene Viewer-Anwendungen haben unterschiedliche, in den meisten Fällen einstellbare Ausspielpuffergrößen. Es ist jedoch wichtig zu bedenken, dass eine Verringerung des Puffers den Jitter erhöht, so dass Sie einen guten Kompromiss zwischen Jitter und tolerierbarer Verzögerung finden müssen.

### 4.3.2 Audio-Puffer

Audiostreaming ist bei der Wiedergabe störungs- und verzögerungsanfälliger als Videostreaming. Ein einzelnes verzögertes Audiopaket erzeugt ein störendes Knacken in der Tonspur, und der Ton muss mit dem Video lippensynchron sein. Aus diesen Gründen benötigen Sie einen großen Ausspielpuffer für Videos mit Audio, was natürlich die Verzögerung erhöht.

### 4.3.3 Dekomprimierung

Die für die Dekomprimierung benötigte Zeit hängt vom verwendeten Kodierungsverfahren ab. Einen großen Ausschlag bei der Dekodierungslatenz gibt die Hardware-Decoder-Unterstützung der Grafikkarte. Die hardwarebasierte Dekodierung ist in der Regel schneller als die softwarebasierte. Im Allgemeinen benötigt die Dekodierung von H.264 mehr Zeit als die von MJPEG. Außerdem hängt die Verzögerung bei der H.264-Dekodierung von dem Profil ab, das Sie für die Kodierungsphase gewählt haben. „Base“ ist am einfachsten zu entschlüsseln, während „Main“ und „High“ mehr Zeit benötigen. Für den von Axis Videoprodukten erzeugten H.264-Videostream muss der Decoder mindestens ein Bild puffern.

### 4.3.4 Anzeigegerät

Auch das Anzeigegerät wirkt sich über seine Übertragungszeit, Bildwiederholrate und Reaktionszeit auf die Verzögerung aus.

Die Übertragungszeit ist die Zeit, die benötigt wird, um die dekodierten Videodaten vom Decoder über das Kabel (z. B. HDMI) an den Monitor zu senden. Sie hängt von der durch das Kabel unterstützten Geschwindigkeit und der Auflösung des Monitors ab. Bei einem FHD-Monitor, der mit einem Standard-HDMI-Kabel angeschlossen ist, bedeutet dies eine Verzögerung von etwa 10 ms.

Außerdem wird die Verzögerung durch die Bildwiederholrate des Anzeigegeräts beeinflusst. Computermonitore haben eine Bildwiederholrate von etwa 17 bis 20 ms, während spezielle Gaming-Monitore eine Bildwiederholrate von 4 bis 5 ms bieten.

Die Reaktionszeit ist die Zeit, die die Pixel im Monitor benötigen, um ihren Wert zu ändern. Sie hängt davon ab, wie groß die Änderung ausfällt, kann aber bei größeren Änderungen einer Verzögerung von 5 bis 20 ms entsprechen.

## 5 Reduzierung der Verzögerung

Die Auslegung eines Systems auf eine geringe Verzögerung erfordert andere Kompromisse. Sie müssen entscheiden, welche Verzögerung tolerierbar ist, und die Videoqualität gegen die Kosten des Sicherheitssystems abwägen. Dieses Kapitel enthält einige einfache Empfehlungen zur Reduzierung der durch die Kamera, das Netzwerk und die Client-Seite verursachten Verzögerung.

### 5.1 Reduzierung der Verzögerung in der Kamera

- **Auflösung.** Wählen Sie nach Möglichkeit eine niedrigere Auflösung. Eine höhere Auflösung bedeutet, dass mehr Daten kodiert werden müssen, was wiederum die Verzögerung erhöhen kann.
- **Verbesserung.** Reduzierung von Bildoptimierungen wie z. B. Rotation, Deinterlacing und Skalierung. Diese können zusätzliche Verzögerungen verursachen.
- **Modus für geringe Verzögerung.** Sie können die Bildverarbeitungszeit Ihres Livestreams durch Einschalten des Low-Latency-Modus in den Einstellungen für die einfache Konfiguration optimieren.

Damit wird die Verzögerung in Ihrem Videostream auf ein Minimum reduziert, was andererseits aber zu einer geringeren Bildqualität als üblich führt.

- **Privatzonenmaskierung.** Überlegen Sie, auf Privatzonenmaskierung zu verzichten, da sie die Verzögerung erhöht.
- **Codierung.** Achten Sie darauf, dass der Encoder den für Ihr System erforderlichen Grad an Kontrolle über die Verzögerung bietet. Die Datenmenge muss auf die Kapazität der Netzwerkinfrastruktur angepasst werden. Wenn das Video über ein Netzwerk mit begrenzter Bandbreite übermittelt wird, wählen Sie H.264 oder H.265 als Kodierungsverfahren. Dies führt zu einer niedrigeren Bitrate aufgrund einer stärkeren Komprimierung. Wählen Sie das Baseline-Profil, wenn das Netzwerk die Verwaltung der Bitrate erlaubt, da das Baseline-Profil einfacher zu kodieren und zu dekodieren ist.
- **Speicherprofil in Zipstream.** Überlegen Sie, auf das Speicherprofil zu verzichten, da es die Verzögerung erhöht.
- **Anzahl der Videostreams.** Begrenzen Sie die Anzahl der Videostreams, die von der Kamera mit verschiedenen Einstellungen aufgenommen werden. Jede einzigartige Konfiguration, wie Auflösung, Bildrate und Komprimierung, erfordert einen eigenen individuellen Kodierungsprozess, der den Prozessor zusätzlich belastet und zu Verzögerungen führt.
- **Bitrate.** Verwenden Sie eine möglichst niedrigere Bitrate. Um die Verzögerung in der Kamera zu verringern, müssen wir die Menge der erzeugten Daten reduzieren. Die Bitrate wird von vielen Faktoren beeinflusst, z. B. von den Lichtverhältnissen, der Art der Szene, der Komprimierungsstufe, der Auflösung, der Bildrate und anderen.
- **Bildrate.** Verwenden Sie eine möglichst hohe Bildrate. Da die Bilder einzeln nacheinander kodiert und dekodiert werden, kommt es in den Puffern zu einer Verzögerung von mindestens einem Bild. Bei höheren Bildraten fallen die in den Puffern verursachten Verzögerungen geringer aus. Bei einem Videostream mit 30 Bildern pro Sekunde dauert die Aufnahme jedes Bildes  $1/30$  einer Sekunde. Dementsprechend ist in den Puffern mit einer Verzögerung von bis zu 33 ms zu rechnen. Bei 25 Bildern pro Sekunde beträgt die Verzögerung bis zu 40 ms.
- **Aufnahmemodus.** Verwenden Sie einen Aufnahmemodus mit einer möglichst geringen Auflösung und einer möglichst hohen Bildrate. Die niedrige Auflösung bedeutet weniger zu verarbeitende Pixel und die hohe Bildrate eine geringere pufferbedingte Verzögerung.
- **Audio.** Überlegen Sie, auf Audio zu verzichten. Audio, das mit dem Video synchronisiert werden muss, erfordert einen größeren Puffer für die Wiedergabe, was zu einer höheren Verzögerung führt.

## 5.2 Reduzierung der Verzögerung im Netzwerk

Viele der Empfehlungen für die Kameraseite zielen darauf ab, die Gesamtmenge der über das Netzwerk gesendeten Daten zu begrenzen. In den meisten Fällen ist ein Netzwerk mit begrenzter Übertragungskapazität der größte Faktor für die Verzögerung. Verfügt das Netzwerk über eine hohe Kapazität, sind viele der oben genannten Empfehlungen hinfällig. Vergewissern Sie sich, dass die Servicequalität (Quality of Service) Ihres Netzwerks gut ist und sämtliche Netzwerk-Sprünge innerhalb des Netzwerks so konfiguriert sind, dass sie Ihrem Video-Bedarf entsprechen. Achten Sie darauf, dass Ihre Bitrate im gesamten Netzwerk garantiert die Übertragung des Datenausgangs Ihrer Kamera unterstützt.

### 5.3 Reduzierung der Verzögerung auf der Client-Seite

Optimierungen auf der Client-Seite haben einen großen Einfluss auf die Verzögerung und in der Regel lässt sich hier eine Menge tun.

**Prozessor und Grafikkarte.** Die CPU spielt eine zentrale Rolle bei der Verzögerung auf der Client-Seite. Verwenden Sie einen leistungsfähigen Prozessor mit ausreichender Kapazität, der den Videostream schnell verarbeiten und gleichzeitig auch andere Anfragen abwickeln kann. Nutzen Sie eine gute Grafikkarte mit der neuesten Software und Unterstützung für die Videodekodierung.

**Viewer/VMS:** Stellen Sie sicher, dass Ihr Viewer keinen unnötig langen Ausspielpuffer nutzt, wägen Sie die Größe aber gegen den Video-Jitter ab.

**Anzeige.** Verwenden Sie einen Bildschirm mit einer möglichst hohen Bildwiederholfrequenz. Für eine angenehme Live-Anzeige (die sich aber nicht auf die Verzögerung auswirkt) stellen Sie die Bildfrequenz auf ein Vielfaches der Bildrate der Kamera ein. Ein Beispiel wären 60 Hz für den Modus mit 30 Bildern/s bzw. 50 Hz für den Modus mit 25 Bildern/s.

# Über Axis Communications

Axis ermöglicht eine intelligente und sichere Welt durch Lösungen zur Verbesserung der Sicherheit und Geschäftsperformance. Als Unternehmen für Netzwerktechnologie und Branchenführer bietet Axis Lösungen in den Bereichen Videosicherheit, Zutrittskontrolle sowie Intercoms und Audiosysteme. Sie werden verstärkt durch intelligente Analyseanwendungen und unterstützt durch gute Schulungen.

Axis beschäftigt rund 4.000 engagierte Mitarbeiter in über 50 Ländern und arbeitet weltweit mit Technologie- und Systemintegrationspartnern zusammen, um den Kunden Lösungen anbieten zu können. Axis wurde 1984 gegründet und der Hauptsitz befindet sich in Lund, Schweden