

# Digitales Diktat unter Citrix

## Technischer Überblick - Grundlagen

Wird voice4branche in einer Terminalserver-Umgebung unter Citrix betrieben, ergeben sich besondere Anforderungen an die beteiligte Infrastruktur.

Die Anwendungen, die Audiodaten aufnehmen bzw. wiedergeben, werden nicht auf dem Clientrechner ausgeführt, an dem auch die entsprechende Aufnahme- bzw. Wiedergabehardware (Mikrofon, Headset, Lautsprecher/Kopfhörer) angeschlossen ist, sondern in einer Anwendersitzung auf einem Applikationsserver (Citrix XenApp Server).

Da der Server, der die Sitzung bereitstellt, über kein physisches Audiogerät verfügt, muss den Anwendungen in der Sitzung ein virtuelles Audiogerät zur Verfügung gestellt werden.

Dieses virtuelle Audiogerät verhält sich aus Sicht der Anwendung wie ein physisches Audiogerät und kann genauso Audiodaten liefern oder zur Wiedergabe entgegennehmen. Unter Citrix wird diese Schicht als **HDX Audio** bezeichnet.

Intern sorgt dieses HDX Audio dafür, dass Audiodaten zur Wiedergabe im virtuellen Audiogerät ggf. komprimiert und über ein Netzwerk an den Client geschickt werden, auf dem die Serversitzung dargestellt wird.

Die dafür zuständige Anwendung (Windows Terminalserver Client bzw. Citrix ICA Client) empfängt die Audiodaten vom Netzwerk, dekomprimiert sie und gibt sie über ein physisch am Client angeschlossenes Audiogerät aus.

Die Aufnahme am Client funktioniert analog, nur ist dort die Clientanwendung für die Aufnahme und Komprimierung zuständig und das virtuelle Audiosystem am Server für den Empfang der Daten übers Netzwerk, die Dekomprimierung und Weitergabe an das jeweilige Aufnahmeprogramm am Server.

## Auswirkungen auf die Infrastruktur

Durch dieses Vorgehen wird – gerade bei vielen gleichzeitig laufenden Diktatsitzungen – das Netzwerk stark belastet.

Je nach verwendeter Citrix-Version und Konfiguration von Servern und Clients kann die Bandbreite für Audio bis zu 1,3 Mbps (Megabit pro Sekunde) erreichen.

Da in aller Regel solche Audiodaten zudem nicht die einzigen Daten sind, die über das Netzwerk übertragen werden müssen (bspw. müssen hunderte MB große DICOM-Dateien kurzfristig bereitgestellt werden), kann es zu Engpässen kommen.

Gleichzeitig lässt sich die verfügbare Bandbreite des Netzwerks nicht beliebig erweitern und es müssen oft auch Endgeräte genutzt werden, deren Netzwerkanbindung nicht besonders üppig ausfällt.

## Ressourcen-Engpässe

Da es sich bei den übertragenen Audiodaten um Echtzeitdaten handelt, spielt die Latenz im Netzwerk eine sehr große Rolle.

Wenn das Netzwerk gerade ausgelastet ist, kann die zeitgerechte Übertragung von Audiopaketen u.U. nicht erfolgen. HDX Audio reagiert normalerweise dadurch auf diese Situation, dass verzögert übermittelte Audiopakete einfach verworfen werden.

Dies kann zu einer merkbaren Verschlechterung der Audioqualität führen, inkl. abgehackten Wörtern oder sogar kompletten Aussetzern.

Bei Transkription durch eine Schreibkraft kann dies zu einem gewissen Grad kompensiert werden (wenngleich es natürlich viel anstrengender ist, eine schlechte Tonqualität zu transkribieren als eine saubere Aufnahme).

Für Spracherkennung gilt dies leider nicht. Solche Aussetzer werden sich unmittelbar in einer erheblich reduzierten Erkennungsgenauigkeit widerspiegeln.

## Alternativen

Für genau diese Herausforderung gibt es für das digitale Diktat mit voice4branche einen eigenen Audiokanal zur Übertragung von Audiodaten.

Dieser virtuelle Audiokanal "P" komprimiert die Audiodaten in ein Format, das hohe Qualität mit geringer Bandbreite kombiniert.

Ein derart komprimierter Audiostrom benötigt trotz sehr guter Sprachqualität lediglich 21 Kbps (Kilobit pro Sekunde) und damit nur ein Drittel dessen, was ein aktuellstes Citrix bei mittlerer Tonqualität benötigt<sup>1</sup>.

Zudem werden Audiopakete so übertragen, dass keine Pakete verworfen werden. Kurzzeitige Netzwerk-Engpässe haben somit keine Auswirkung auf die Audioqualität.

## Installation

Für diesen virtuellen Audiokanal muss serverseitig das 4voice Treiberpaket installiert werden. Dies enthält u.a. die dafür notwendigen Server-Komponenten.

Daneben sind die Philips Citrix Extensions erforderlich – je nach Version server- und clientseitig oder nur am Client.

Bei Windows-basierten Clients lassen sich diese Komponenten einfach installieren.

Linux-basierte ThinClients (z.B. IGEL UD-LX) bringen zwar oft die entsprechend an Linux angepassten Clientkomponenten in ihrer Firmware mit, aber das gilt leider nicht für alle verfügbaren Hersteller und Geräte.

Daher muss jeweils bezogen auf konkrete ThinClient-Modelle die Verfügbarkeit der Clientkomponenten geklärt werden.