



UNIFY<sup>2</sup>

*Unified Communications, Accelerated*

---

Sind Voice VLANs relevant für die Sicherheit in Unified Communications Umgebungen?

Unify<sup>2</sup> Inc.

19. Januar 2010

[www.unifysquare.com](http://www.unifysquare.com)

## Inhalt

1	Einleitung.....	3
2	VLANs Überblick .....	3
3	VLANs und VoIP/UC Umgebungen .....	6
4	Auswirkungen der Benutzung von VLANs als Sicherheitsmaßnahme für VoIP/UC.....	8
5	Wie begegnet nun OCS diesem Risiko? .....	8
6	Schlussfolgerung.....	9
	Referenzen .....	10

## 1 Einleitung

Die Verwendung von Virtual Local Area Networks (VLANs) wurde bislang in IP-Telefonieumgebungen als eine Grundvoraussetzung an die Netzwerkinfrastruktur angesehen. Dabei sollten VLANs neben anderen Vorteilen vor allem einen zuverlässigen Einsatz und hohe Sicherheit garantieren.

Weiterentwicklungen im Voice-over-IP- (VoIP) und Unified Communications- (UC-) Umfeld wie bspw. Microsoft's Office Communications Server 2007 R2 sehen die Verwendung von VLANs nicht länger als eine Grundvoraussetzung an die Netzwerkinfrastruktur an. Hinzukommt, dass heute Open-Source Programme frei herunterladbar im Internet existieren, die unverschlüsselten Sprach- und Videodatenverkehr über VLANs mitschneiden können. Hierdurch werden sogenannte "Man-in-the-Middle" Angriffe auf Unified Communications-Medienströme ermöglicht. Diese Entwicklungen stellen zunehmend den Sicherheitsaspekt von VLANs in Frage, welcher bislang als Hauptgrund für die Verwendung von dedizierten Sprach-VLANs (Voice VLANs) in IP-Telefonieumgebungen angesehen wurde. Dies hat zur Folge, dass Voice VLANs nicht mehr uneingeschränkt als Sicher gegen Mitschneiden oder Manipulationen angesehen werden können, sofern in den Voice LANs unverschlüsselte Medienströme (Sprache, Video oder auch Desktop Sharing) übertragen werden.

Hacking Software wie bspw. [UCSniff](#) oder [VideoJak](#) erlauben einen einfachen Zugriff auf Sprach- und Videomedienströme. Es lässt sich sogar dediziert auswählen, welche spezifischen Medienströme überwacht oder manipuliert werden sollen. Folglich sind viele, wenn nicht sogar die meisten Lösungen, die Sprach- und Videodatenverkehr unverschlüsselt in Sprach-VLANs übertragen, potenzielles Ziel für solche Angriffe geworden.

All die oben angeführten Entwicklungen zusammengenommen stellen das bislang manifestierte Grundkonzept von Sprach-VLANs als Sicherheitsmethode für VoIP-Lösungen in Frage.

Diese Ausarbeitung erlaubt dem Leser, generelle und historische Gründe für VLANs im Allgemeinen und Sprach-VLANs im Besonderen zu verstehen und zeigt auf, dass VLANs nicht wirklich zur Sicherheit im Sinne einer Absicherung der Ende-zu-Ende Medienströme gegen Mitschneiden, Überwachung und Manipulation in VoIP- oder UC-Umgebungen beitragen. Es wird ferner aufgezeigt, wie insbesondere standardmäßig implementierte IP-Telefonielösungen zur Absicherung oftmals zusätzlicher Zugangsrestriktions- oder anderer Netzwerksicherheitsfunktionen bedürfen, obwohl diese mittels bekannter Verfahren und Hilfsprogrammen umgangen werden können, sofern die Implementierungen auf Basis von VLANs erfolgte. Ebenfalls wird aufgezeigt, wie der Zugriff auf unverschlüsselte Sprach-VLANs technisch realisiert wird. Zum Schluss wird erläutert, wie Microsoft's Office Communications Server (OCS) 2007 R2 bereits bei einer Standardimplementierung Sicherheit gegen Mitschneiden, Überwachung oder Manipulation von Medien- und Signalisierungsströmen unter Verwendung von Industriestandards bietet.

## 2 VLANs Überblick

VLANs wurden im Rahmen der Entwicklung des Ethernet-Verkehrsmanagements entwickelt. Das Ethernet, welches im IEEE Standard 802.3 definiert ist und auf dem OSI-Layer 2 operiert, benutzt konzeptionell eine sehr einfache Methode um mehreren Geräten das Senden von Daten über eine gemeinsame physikalische

Verbindung zu ermöglichen. Dabei existiert keinerlei Kontrolle darüber, welches Gerät zu welchem Zeitpunkt Daten über das gemeinsame Medium (hier: die LAN-Infrastruktur) senden darf. Dieses Verfahren wird Carrier Sense Multiple Access with Collision Detection (CSMA/CD) genannt, ist immanenter Bestandteil nahezu aller modernen LAN Umgebungen und verdrängte andere Technologien wie bspw. Token Ring oder ATM.

Jedes Gerät darf zu jedem Zeitpunkt Daten über das gemeinsame Medium (hier: Netzwerk) schicken, sofern nicht just in dem Moment des Sendens erkannt wird, dass bereits ein anderes Datenpaket auf dem Netzwerk übertragen wird. Sollten dennoch zwei Geräte gleichzeitig Daten über das Netzwerk senden, kommt es zu einer Kollision der Datenpakete, welche von den Geräten erkannt wird. Anschließend werden beide Geräte eine zufällig bestimmte Zeitdauer abwarten, bevor sie ihre Daten erneut über das Netzwerk zu übertragen versuchen. Da jedes Gerät theoretisch eine andere zufällige Zeitdauer vor der Neuaussendung wartet, wird bei der nächsten Versendung der Datenpakete eine wiederholte Kollision unwahrscheinlich.

Dieses Verfahren skaliert natürlich nur bis zu einem gewissen Grade. Je höher die Anzahl der angeschlossenen Geräte auf einem Netzwerk ist, desto öfter wird es auch zu Kollisionen von Datenpaketen auf dem Netzwerk kommen. Im Extremfall kann es sogar bis zu dem Zustand führen, dass keine von den Geräten gewählte zufällige Wartezeit vor einer Neuaussendung zu einer Verhinderung von Kollisionen führt. Somit kann kein Gerät mehr Daten über das Netzwerk senden, da stets ein anderes Gerät ebenfalls zum gleichen Zeitpunkt Daten senden möchte. Zusätzlich existieren auch noch sogenannte Broadcast-Nachrichten, welche zwar nur für einen Teil der angeschlossenen Geräte relevant sind, aber von allen am Netzwerk angeschlossenen Geräten empfangen und ggfls. sogar beantwortet werden müssen. Steigt die Anzahl der angeschlossenen Geräte auf einem Netzwerk, wird auch die Anzahl der Broadcast-Nachrichten zunehmen und somit auch die Menge der zu verarbeitenden Nachrichten durch jedes angeschlossene Gerät. Dies allein nur deshalb, weil ein Gerät als Empfänger einer Broadcast-Nachricht potenziell in Frage kommt. Darüber hinaus muss jedes Gerät natürlich auch noch diejenigen Nachrichten verarbeiten, die bereits ursprünglich für das dedizierte Gerät bestimmt waren. Wird der oben beschriebene Verkehr zusammengenommen, kann es dazu führen, dass die Geräte häufiger nicht mehr in der Lage sind, in effektiver Weise Datenpakete resp. Nachrichten zu senden oder zu empfangen.

Teilweise als Antwort auf das geschilderte Problem der Netzwerkkollisionen wurde sog. Ethernet Switches entwickelt, welche mit ihren Funktionalitäten dieser Problematik entgegenwirken. Ein moderner Switch bietet heute vielfältige Möglichkeiten, um Broadcast- und auch dedizierte Nachrichten derart zu verwalten, dass es nicht mehr zu einer Überlastung des Netzwerks kommt. Darüber hinaus wird die Anzahl der Geräte vermindert, die Nachrichten anderer Geräte mitverarbeiten können und müssen.

Zunächst einmal unterbricht ein Switch die Verbindungen zwischen allen Geräten und sendet nur Broadcast- und dedizierte Nachrichten zu den angeschlossenen Geräten, für die die jeweiligen Nachrichten auch wirklich bestimmt sind. Ein Gerät ist nicht mehr in der Lage, die Nachrichten zwischen zwei anderen angeschlossenen Geräten zu empfangen und ggfls. zu verarbeiten. Es sei denn, dass alle Geräte am selben Switchport angeschlossen sind, der Switch mit Daten überschwemmt wurde oder das ARP (Address Resolution Protocol) manipuliert wurde.

Eine weitere Eigenschaft eines Switches ist, dass er dem Administrator erlaubt, ein LAN in kleinere Segmente zu unterteilen, wobei der Verkehr zwischen diesen Segmenten durch den Switch unterbunden wird. Dies ist das klassische Virtual LAN (VLAN). Virtueller, weil obwohl die Geräte immer noch auf die gleiche Art und Weise an demselben physikalischen Switch angeschlossen sind, können sie keine Daten direkt mehr zwischen den einzelnen VLANs senden. Wurde ein LAN in verschiedene VLANs segmentiert, benötigt es der Funktionalität

eines Netzwerkelementes, das auf einer höheren Netzwerkschicht arbeitet, um Nachrichten in andere VLANs zu senden. Typischerweise ist dies ein Router, der auf dem Layer 3 des OSI-Modells operiert.

Schließlich kann ein Administrator noch mehrere Switche zusammenschließen, wobei angeschlossene Geräte des gleichen VLANs auf unterschiedlichen Switchen weiterhin miteinander kommunizieren können. Diese Eigenschaft ist vielleicht eine der interessantesten, weil es dem "LAN" Konzept eine weitere Dimension verleiht. Es werden komplexe Netzwerkarchitekturen ermöglicht, die von den Switchen erfordern, Datenpakete von anderen Switchen zu vertrauen, welche von diesen für bestimmte Ziel-VLANs markiert wurden.

Der zuletzt genannte Punkt bedarf einer besonderen Betrachtung, da er in direktem Zusammenhang mit Sprach-VLANs in IP-Telefonie- oder UC-Umgebungen steht.

VLANs sind nützlich für ein zentralisiertes Netzwerkmanagement und um physikalisch voneinander separierten aber logisch zueinander gehörigen Geräten zu ermöglichen, innerhalb einer IP-Broadcast-Domäne oder eines Subnetzes miteinander zu kommunizieren. Betrachtet man VLANs nur unter Auslastungsgesichtspunkten, erlauben sie Unternehmen die Anzahl der benötigten Switche und der benötigten Ports für die Inter-Switch-Kommunikation zu reduzieren.

Etwas weniger gewichtig sind die Eigenschaften, dass VLANs als Hilfsmittel zur Paketpriorisierung und zum Bandbreitenmanagement an dedizierten Switchports verwendet werden können.

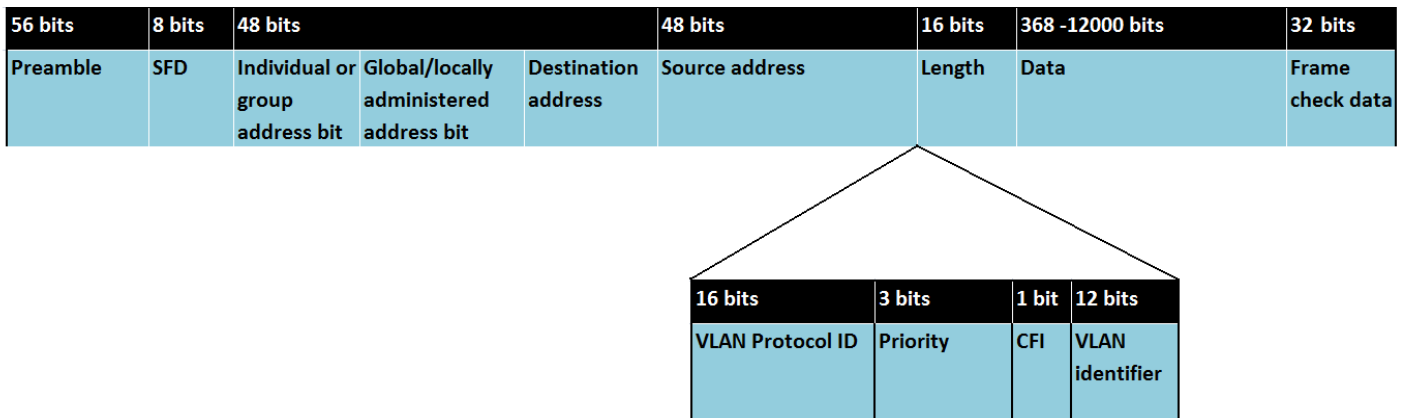
VLANs bieten eine Vielzahl von Vorteilen:

- Reduzierung der Anzahl benötigter Switche in einem Netzwerk
- Reduktion des administrativen Overheads für Netzwerkmanagement
- Vereinfachte Medium Access Control (MAC) für den Zugang auf ein bestimmtes Netzwerk
- Vereinfachte Fehlersuche in komplexen Netzwerkkumgebungen

Was VLANs hingegen NICHT bieten, sind Sicherheitsmechanismen für Sprachverkehr.

Um zu verstehen, wie Sicherheit erzielt wird und welche Implikationen durch die Aktivierung von Sicherheitsmechanismen entstehen, muss zunächst verstanden werden, wie Datenverkehr in einem VLAN markiert wird. Im Folgenden ist der IEEE 802.3 Standard für VLAN Rahmen auf einem Ethernet abgebildet:

Bild 1: IEEE 802.3 Rahmen Format



Dem Standard folgend kann ein Ethernet Rahmen *optional* ein [IEEE 802.1Q](#) Tag beinhalten (in Bild 1 unten), welches Aussage über die [VLAN](#)-Zugehörigkeit und die [IEEE 802.1p](#) Priorität ([quality of service](#)) gibt. Die Einbindung dieses Tags ist definiert in der [IEEE 802.3ac](#) Spezifikation und vergrößert die Rahmengröße um 4 Bytes auf 1522 Bytes.

Dies wird erreicht, indem der [IEEE 802.1Q](#) Tag mit einer Größe von 4 Bytes zwischen dem Senderadressfeld (Source Address Field) und dem Längen/Typ-Feld (length/type) im Header des Ethernet-Rahmens eingefügt wird. Neben anderen Identifikatoren beinhaltet der Header auch die Information über das aussendende VLAN (Originating VLAN) und ein 3-bit Priorisierungsfeld, das für die Implementierung der Priorisierung nach 802.1p auf dem Ethernet benötigt wird.

Die ersten zwei Bytes des IEEE 802.1Q Tags sind der Tag Protokoll Identifikator (Tag Protocol Identifier (TPID)) mit dem Wert 0x8100. Der TPID bei markierten Ethernet-Rahmen wird an demselben Ort platziert wie auch bei nicht markierten Ethernet-Rahmen. Der Wert 0x8100 indiziert, dass der Rahmen markiert wurde und der eigentliche EtherType/Length-Identifikator befindet sich nach dem Q-Tag. Der TPID wird gefolgt von der 2 Byte großen Tag Control Information (TCI) (der IEEE 802.1p Priorität ([quality of service](#)) und [VLAN ID](#)).

### 3 VLANs und VoIP/UC Umgebungen

VLANs wurden historisch als eine empfohlene Investition in die Netzwerkinfrastruktur für IP-Telefonieumgebungen angesehen. In einer typischen Konfiguration werden der Unternehmens-Telefonieverkehr (Signalisierung und Medienströme) getrennt vom übrigen Unternehmensdatenverkehr (PCs, Applikationen) über separate VLANs – dem "Sprach-VLAN" und dem "Daten-VLAN" – über ein gemeinsames physikalisches Netzwerk (Switches, Router, Kabel) übertragen. Netzwerkverkehr, der von einem IP Telefon gesendet wird, erhält die Markierung des Sprach-VLANs. Dies erlaubt der Netzwerkinfrastruktur die einfache Unterscheidung zwischen zwei Typen von Paketen. Ein solches Design soll vor allem die folgenden Vorteile bieten:

1. Automatische Priorisierung der Sprachdaten: Durch die automatische Markierung aller Pakete des Sprach-VLANs mit einem höheren QoS-Wert gegenüber den Daten-VLAN-Paketen können Switches

und Router Sprachpakete mit höherer Priorität behandeln als Datenpakete, wodurch schließlich Latenz und Varianz des Sprachdatenverkehrs verringert wird.

2. Einfachere Implementierung: Durch eine übereinstimmende Konfiguration der IP-Telefonieendgeräte und der Switch/Router-Infrastruktur kann das Unternehmen eine QoS-Implementierung sicherstellen, die keinerlei Maßnahmen vom Endbenutzer bedarf und eine robuste Gesamtlösung bietet.
3. Sicherheit: Durch die Konfiguration eines separaten VLANs sollte durch das Hinzufügen einer zusätzlichen virtuellen Trennschicht zwischen dem Sprachdaten- und dem Datenverkehr eine höhere Sicherheit für den Sprachdatenverkehr erzielt werden, da der Datenverkehr typischerweise häufiger Angriffen durch Mitschnitten ausgesetzt ist.

Bei der Untersuchung der Auswirkungen von VLANs auf Unified Communications ist es zunächst wichtig zu verstehen, dass die Implementierung einer UC-Lösung sich wesentlich von der einer IP-Telefonielösung unterscheidet. Eine IP-Telefonielösung verwendet hauptsächlich IP-Telefone als Endpunkte. Diese Geräte eignen sich für die Verwendung von VLANs, weil ausschließlich Sprachdatenverkehr von dem jeweiligen Endgerät ausgesendet wird. Dies ermöglicht die Segmentierung des Sprachdatenverkehrs in ein dediziertes Sprach-VLAN, weil das Endgerät über eine eindeutige IP- und MAC-Adresse verfügt.

Unified Communications geht weit über die Verwendung von IP-Telefonen hinaus und erlaubt die Verwendung multimodaler PC-Desktop-Software, die multimediale Zusammenarbeit ermöglicht. Im Laufe der Zeit werden die Funktionalitäten, die derzeit noch eines dedizierten Clients bedürfen, in die Standard-Geschäftsanwendungen integriert werden und untrennbar direkt aus diesen nutzbar sein. Dies führt dazu, dass die traditionelle Trennung in Unternehmens-Sprachdaten- und Datenverkehr nicht mehr möglich sein wird. Darüber hinaus entstehen technische Komplexitäten, wenn versucht wird, verschiedene Datenströme auf mehrere VLANs zu verteilen, während der Desktop-PC üblicherweise an einem einzigen LAN angebunden ist.

Eine Alternative zu VLANs ist die Verwendung des Differentiated Services (DiffServ) Quality of Service (QoS) Protokolls. Dieses Verfahren ermöglicht Desktop Applikationen ihren Datenverkehr eindeutig zu markieren. Dabei kann die Switch-Infrastruktur diese Pakete vom übrigen Netzwerkverkehr unterscheiden und ggfls. mit höherer Priorität verarbeiten. Dies führt zu einem ähnlichem Verhalten der logischen Unterscheidung von Paketen wie es auch bei den VLANs der Fall ist, jedoch ohne die Notwendigkeit einer logischen Netzwerkseparation (wie in Kapitel 2 erwähnt, bietet die logische Netzwerksegmentierung keine zusätzlichen Sicherheitsvorteile).

Office Communications Server (OCS) unterstützt DiffServ Markierung auf dem Server, dem Desktop-PC und dem IP-Telefon-Endpunkten. Dies ermöglicht die eindeutige Markierung vom Sprachdaten- und dem Videodatenverkehr mit zwei Prioritätsinformationen, dass die Switch-Infrastruktur diesen Netzwerkverkehr dem restlichen "Best Effort" Verkehr bevorzugen kann. Die Absicherung der Sprachdaten- und Videodatenpakete ist losgelöst von der Markierung der Pakete für eine bevorzugte Übertragung. OCS bietet standardmäßig eine sichere Ende-zu-Ende-Lösung, deren Funktionsweise im Detail in Kapitel 5 erläutert wird.

Wie in den nachfolgenden Kapiteln aufgezeigt wird, bieten VLANs in der Praxis keine zusätzliche Sicherheit gegenüber normalen LANs. Darüber hinaus vereinfachen VLANs nicht mehr die Implementierung von UC-Umgebungen, weil UC Endgeräte i.d.R. wesentlich mobilere Geräte als IP-Telefone sind und die Mobilität von Endgeräten ein kritischer Bestandteil einer erfolgreichen VoIP/UC-Implementierung für ein Unternehmen ist.

## 4 Auswirkungen der Benutzung von VLANs als Sicherheitsmaßnahme für VoIP/UC

Neue Open-Source Programme machen es einfach, Sprachdaten- und Videodatenverkehr auf VLANs mitzuschneiden.

Erst kürzlich wurde auf der DefCon 17 Hacking Conference in Juli/August 2009 von den Entwicklern des VoIP Sniffers UCSniff - Jason Ostrom and Arjun Sambamoorthy - eine neue Version 3.0 des UCSniff-Tools vorgestellt, welches eine klassische "Man-in-the-Middle"-Attacke auf einem Unternehmens-VLAN zum Mitschneiden von Sprachverbindungen ermöglicht. UCSniff verwendet die "Ettercap-Suite" für "Man-in-the-Middle"-Attacken auf LANs. Wenn das Programm Zugriff auf einen Ethernet-Switch eines Unternehmens erhält, welches Sprach-VLANs verwendet, erkennt das Programm automatisch die VLAN IDs diverser Hersteller (Cisco, Avaya, Nortel) und führt dann eine ARP Manipulationsattacke durch, um einen Zielpunkt für eine VoIP-Kommunikation zu simulieren.

UCSniff optimiert das Mitschneiden von Gesprächen von dedizierten Benutzern. Ziele können durch die Nebenstelle oder durch Namenswahl ("Dial-by-name") ausgewählt werden, was das Mithören /-schneiden von Gesprächen einer bestimmten Person (z.B. dem CEO) einfach gestaltet. Das Mithören /-schneiden kann weiter optimiert werden, in dem das Tool es zulässt, dass nur Gespräche zu bestimmten Zielpersonen (z.B. zwischen dem CEO und dem CFO) mitgehört/mitgeschnitten werden.

Die Entwickler von UCSniff haben Video als weitere Funktionalität hinzugefügt. Diese Funktionalität heisst VideoJak. VideoJak kann willkürliche Videosequenzen in einen Videodatenverkehr einspeisen, wodurch bspw. Videoüberwachungsprogramme kompromittiert werden können. In einer Live-Demonstration haben die Entwickler gezeigt, wie die Juwelen eines Museums (simuliert durch eine Wasserflasche) trotz Videoüberwachung gestohlen werden können. Dabei wurde von dem Tool zunächst für 20 Sekunden die unberührte Flasche aufgezeichnet und dann diese Sequenz wiederholt abgespielt. Währenddessen hatte der "Dieb" genug Zeit, die Flasche an sich zu nehmen, unbeobachtet von dem Sicherheitspersonal.

## 5 Wie begegnet nun OCS diesem Risiko?

Bevor nun weitere Details erläutert werden, wie OCS eine sichere Implementierung seiner UC-Lösung über ein Netzwerk erreicht, ist es wichtig zu verstehen, dass *OCS standardmäßig ("out-of-the-box") sicher* ist. Viele andere IP-Telefonieumgebungen bieten Sicherheit als eine Zusatzfunktionalität an, die für jeden Endpunkt zusätzlich konfiguriert werden muss, nachdem die Lösung zuvor bereits implementiert wurde. Dies kann zu einer unsicheren Implementierung führen, wenn die IT-Abteilung die Zusatzfunktionalität "Sicherheit" nicht konfiguriert oder Fehler bei der Konfiguration macht. OCS geht den umgekehrten Weg, weil er alle Kommunikationsmodalitäten standardmäßig sicher anbietet. Das Grundprinzip, dass OCS Kommunikation standardmäßig sicher ist, hilft OCS Implementierungen dabei, stets sicher zu sein. Unabhängig davon, wie auch immer das darunterliegende Netzwerk aufgebaut ist.

Um zu verstehen, wie OCS eine sichere Kommunikation ermöglicht, muss ein kurzer Blick auf die verwendeten unterliegenden Protokolle in einer IP-basierten Sprachkommunikation geworfen werden. Grundsätzlich ist jede IP-Kommunikation (Sprache, Video, Webkonferenzen, Desktop Sharing ...) ein Mix aus

zwei separaten Protokollen: Zum einen SIP (Session Initiation Protocol) als Signalisierungsprotokoll und zum anderen RTP (Real Time Protocol) als Medienübertragungsprotokoll. Bei beiden Protokollen handelt es sich um von der IETF verbreitete Standards für die IP Kommunikation. Das SIP Protokoll wird dabei für die Authentifizierung, die Autorisierung, die Behandlung von Angebots- und Antwort-Anfragen und die Aushandlung der Verbindungsführung und der Kommunikationscodecs zuständig. Das RTP-Protokoll kümmert sich dabei um den eigentlichen Medienstrom (Payload).

Bspw. wird der SIP Signalisierungskanal für die Authentifizierung des Clients und die Registrierung am OCS Server sowie die Versorgung des Clients mit konfigurationsrelevanten Informationen vom Server (wie, dass der Benutzer für Sprachtelefonie freigeschaltet wurde) verwendet. Wenn der Benutzer nun einen Anruf tätigt, wird der gleiche SIP Signalisierungskanal für das Austauschen der Angebot-/Antwortenabfolge zwischen den beteiligten SIP-Endpunkten zur Erstellung der Sprachverbindung verwendet. Nachdem der Übertragungsweg und Codec zwischen den an der Kommunikation beteiligten Endpunkten ausgehandelt wurden, beginnen die beiden Kommunikationspartner mit dem Senden von RTP Medienpaketen unter Verwendung des zuvor ausgehandelten Codecs.

OCS ist eine sichere Ende-zu-Ende Lösung, da Signalisierungs- und Medienprotokolle stets verschlüsselt werden, was die Notwendigkeit der Implementierung von Sicherheitsmechanismen in den darunterliegenden Netzwerkschichten eliminiert. Dies funktioniert folgendermaßen:

- Zwischen OCS Clients und Servern wird die SIP Kommunikation unter Verwendung von TLS (Transport Layer Security) geschützt. Während der TLS Aushandlung validiert zunächst der Client das vom Server erhaltene Zertifikat, ob der darin enthaltene Domänenname des Servers von einer zertifizierten Zertifikatsstelle signiert wurde. Dann wird ein sicherer Kanal zwischen dem Client und dem Server etabliert, der darin alle SIP Kommunikation mit 128 bit symmetrisch verschlüsselt.
- Für den Fall, dass Signalisierungsinformationen zwischen zwei OCS Servern ausgetauscht werden, wird die SIP Kommunikation auf ähnliche Weise gesichert. Beide Server validieren das Zertifikat des anderen Servers und überprüfen, ob der andere Server zusätzlich in einer "trusted list" als bekannte OCS Rolle enthalten ist.

Wird ein Telefongespräch nun aufgebaut, wird zunächst ein Schlüssel zur Verschlüsselung des Medienstromes über die gesicherte SIP Verbindung aufgebaut und anschließend der Medienstrom über SRTP (Secure Real Time Protocol) verschlüsselt. Dieses Protokoll verwendet ebenfalls einen 128 bit Schlüssel zur Sicherung des Medienstromes.

Indem sowohl das SIP Signalisierungsprotokoll durch TLS wie auch der RTP Medienstrom durch SRTP verschlüsselt werden, wird es einem Angreifer unmöglich gemacht, eine OCS-IP-Kommunikation mitzuschneiden oder auch anderweitig zu manipulieren. Dies sichert die OCS Kommunikation standardmäßig "out-of-the-box" unabhängig von der darunterliegenden Netzwerkinfrastruktur, sei es nun ein Unternehmens-LAN, ein VLAN oder auch das Internet.

## 6 Schlussfolgerung

Unternehmen verwenden bislang häufig Sprach-VLANs für die Übertragung von IP-Telefonieverbindungen. Ein Hauptgrund dafür war die Sicherung des Sprachdatenverkehrs. Jedoch neuartige Sniffing- und VLAN-hopping tools eliminieren den vermeintlichen Sicherheitsvorteil von VLANs. Kunden mit IP-Telefonielösungen, die eine Sprachdatenverschlüsselung nur systemweit oder garnicht zulassen und darüber hinaus ältere IP-Telefone betreiben müssen, welche nicht für Verschlüsselung upgedated werden können, sind der Gefahr ausgesetzt, dass durch "Man-in-the-Middle"-Attacken die Sicherheit der Kommunikation nicht mehr gewährleistet werden kann und Gespräche mitgeschnitten werden können. Anstatt Sprach-VLANs zur Sicherung von IP-basierten Sprachdatenverbindungen zu verwenden, etabliert sich zunehmend die Forderung nach Ende-zu-Ende-Verschlüsselung als eine Grundvoraussetzung für eine Voice-over-IP oder eine Unified Communications Umgebung. Durch die standardmäßige "out-of-the-box" Unterstützung einer Ende-zu-Ende-Verschlüsselung bietet der Microsoft Office Communications Server eine ideales Fundament für den Aufbau einer sicheren Unified Communications Lösung, und zwar unabhängig von dem darunterliegenden Netzwerk.

## Referenzen

Cisco Unified Communications SRND Based on Cisco Unified Communications Manager 6.x  
[http://www.cisco.com/en/US/docs/voice\\_ip\\_comm/cucm/srnd/6x/security.html](http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/6x/security.html)

Cisco Unified Communications Manager Security Guide, Release 7.0(1)  
[http://www.cisco.com/en/US/docs/voice\\_ip\\_comm/cucm/security/7\\_0\\_1/secugd/secu\\_ph.html](http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/security/7_0_1/secugd/secu_ph.html)

Artikel "heise Security": Angriff auf Audio- und Videokonferenzen wird zum Kinderspiel  
<http://www.heise.de/security/meldung/Angriff-auf-Audio-und-Videokonferenzen-wird-zum-Kinderspiel-749461.html>

English version:  
<http://www.h-online.com/security/news/item/DEFCON-Attack-on-audio-and-video-conferencing-made-easy-742777.html>