UNIVERSITÀ DEGLI STUDI DI SALERNO
FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
CORSO DI LAUREA SPECIALISTICA IN INFORMATICA

SICUREZZA SU RETI II

# VOIP SECURITY

STUDENTI:
ANGELO REPPUCCI
ANTONIO PICARDI

DOCENTE:
PROF. ALFREDO DE SANTIS

ANNO ACCADEMICO 2007/2008

# Indice

# Introduzione

La rapida adozione della telefonia IP ha portato a una crescita delle problematiche legate alle vulnerabilità e a un aumento proporzionale delle minacce alla sicurezza. Attualmente, ci si chiede quindi se l'integrità e la sicurezza delle applicazioni su reti convergenti siano davvero in pericolo. Molte delle problematiche che riguardano la sicurezza della telefonia sono simili a quelle già affrontate dalla tecnologia tradizionale e dai sistemi basati su IP. Nonostante ciò, la telefonia IP introduce nuove aree di vulnerabilità che se non vengono gestite correttamente possono essere soggette più facilmente a minacce. La maggior parte degli utilizzi VoIP (Voice over IP) sono stati realizzati dando priorità alle prestazioni e alla qualità del servizio delle reti convergenti. Erroneamente, la sicurezza viene spesso considerata come un fattore a posteriori. Dal momento che i servizi di telefonia aziendali critici operano adesso su una rete multiservizio che, se trascurata, può mettere in pericolo la stabilità e l'integrità dei servizi sopra citati.

Un report del Gennaio 2005 realizzato dal National Institute of Standards and Technology affermava: "I sistemi VoIP potrebbero essere più vulnerabili rispetto ai sistemi di telefonia convenzionali e questo, in parte, è dovuto al fatto che sono collegati a una rete dati e, di conseguenza, risultano più deboli e soggetti ad attacchi." Fortunatamente, la maggior parte delle reti IP hanno già installate funzionalità per la sicurezza e l'esperienza maturata nell'ambito della voce tradizionale può aiutare a risolvere problematiche comuni. Sfruttando le funzionalità di entrambi gli ambiti, l'IPT (IP Telephony) può diventare sicuro, e ancora più sicuro, rispetto ai sistemi di telefonia tradizionali. Le minacce all'IPT possono essere così catalogate:

1. Appropriazione indebita del servizio – Frodi sui prezzi attraverso utilizzi non autorizzati delle risorse telefoniche.

2. Negazione del servizio (Denial of Service) – Monopolizzazione delle

risorse, in maniera accidentale o volontaria, con lo scopo di impedire le normali operazioni.

3. Intercettazione di comunicazioni – attacco alla privacy attraverso l'ascolto delle conversazioni telefoniche o accessi non autorizzati a servizi di messaggistica, come la voicemail.

Nel corso della trattazione prenderemo in considerazione soltanto le ultime due minacce citate prima, dando invece un accenno in questa introduzione di come risolvere la prima.

L'appropriazione indebita del servizio è una problematica legata sia alla definizione delle policy e al rilascio di un applicativo, sia agli aspetti tecnologici. La tecnologia può essere considerata ben indirizzata se è in grado di identificare gli eventi di appropriazione indebita del servizio e garantire un rafforzamento delle policy per prevenire questi eventi. Attualmente, sono disponibili strumenti in grado di individuare le anomalie nello schema delle chiamate, possibili frodi e abusi telefonici. L'attenuazione di questo genere di minaccia negli utilizzi VoIP aziendali è possibile attraverso l'applicazione delle best practice e il potenziamento delle policy che comprende soggetti come:

- Password più solide
- Accesso amministrativo limitato ai gateway voce e soft-PBX(Private Branch Exchange)
- Implementazione di spazi adeguati per la ricerca delle chiamate e restrizione delle chiamate
- Autenticazione degli end-point e degli utenti al sistema
- Configurazione appropriata dei piani telefonici
- Strumenti per monitorare lo schema delle chiamate

Nonostante molte delle vulnerabilità menzionate siano comuni sia alle reti tradizionali sia a quelle IP, l'identificazione e la mitigazione di queste minacce può differire considerevolmente specialmente nell'ambito degli attacchi Denial of Service (DoS).

Dal momento che le attuali applicazioni sono basate su protocolli standard aperti, e quindi subito disponibili, i protocolli di comunicazione vengono esposti al libero arbitrio di chiunque abbia accesso a Internet o a una library pubblica. Questo non consente la "sicurezza attraverso

la segretezza", comune nel mondo della telefonia tradizionale. L'architettura base per le reti IP le rende vulnerabili agli attacchi DoS a meno che vengano configurati specificamente per identificare e limitare i loro effetti. Gli attacchi DoS possono essere gestiti a due livelli: rete e applicazione/sistema operativo.

A livello di rete si può operare come segue:

- Limitare il traffico ai servizi strategici alle porte IP necessarie per azionare, bloccare tutto il resto del traffico verso questi dispositivi
- Segmentare le reti utilizzando Virtual LAN (VLAN) all'interno delle reti logiche che hanno funzioni specifiche (voce, dati, dispositivi, gestione, ecc.)
- Traffico a tasso limitato per prevenire attacchi che deteriorano le risorse di rete, ostacolando le normali operazioni
- Implementare IPS(Intrusion Protection System) e firewall di rete per proteggere i principali servizi sulla rete, tra cui il software PBX, i gateway voce e le applicazioni server (Voice mail, Telephone Management Server, ecc.)
- Abilitare configurazioni di rete per la sicurezza per garantire corrette operazioni di DHCP, DNS e altri servizi di rete.

Mentre a livello di applicazione/sistema operativo si può operare come segue:

- Garantire che le applicazioni siano aggiornate in linea con i più recenti livelli di sicurezza
- Proteggere i server attraverso l'utilizzo di IPS host e software antivirus
- Garantire che solo i servizi essenziali siano disponibili per tutti i sistemi operativi
- Disabilitare i servizi inutili sugli endpoints, come , HTTP, ecc.

Infine, la terza minaccia, che viene molto spesso trascurata, è relativa alle intercettazioni telefoniche e alla violazione della privacy personale. Si può manifestare in diverse forme ma le più comuni sono:

- Compromettere la Voice Mail e i server di messaggistica unificata per ottenere l'accesso ad altri messaggi degli utenti

- Intercettazione dei pacchetti voce sulla rete con lo scopo di riassemblarli e introdursi all'interno di altre conversazioni
- Ottenere l'accesso fisico a dispositivi analoghi o gateway voce per attaccare gli strumenti di intercettazione telefonica

I sistemi di telefonia IP possono essere rafforzati contro queste forme di intercettazioni telefoniche proteggendo le applicazioni nello stesso modo indicato sopra, rispettivamente abilitando la crittografia del traffico voce sulla rete IP e garantendo la sicurezza fisica. Molte delle minacce alla sicurezza della telefonia IP aziendale non riguardano solo la telefonia ma anche le applicazioni IP di rete. L'implementazione delle metodologie per la sicurezza per la telefonia IP migliorerà prima di tutto la sicurezza di molte delle altre applicazioni di rete.

Nel primo capitolo daremo un'introduzione generale sulle reti di telecomunicazioni. Nel secondo capitolo inizieremo a trattare la tecnologia VoIP dandone una definizione generale e illustrando le motivazioni per le quali conviene utilizzare questa tecnologia rispetto a quella tradizionale, mentre nel terzo capitolo forniremo una trattazione completa sui protocolli VoIP maggiormente utilizzati oggi, passando prima per un'illustrazione del funzionamento della telefonia tradizionale. Nel quarto capitolo parleremo di quelle che risultano essere le problematiche di sicurezza di un sistema VoIP. Nel quinto capitolo parleremo del Denial of Service, il tipo di attacco che risulta avere maggiore successo nei confronti delle reti VoIP in quanto va a compromettere un requisito fondamentale di questa tecnologia: la Quality of Service. Nel sesto capitolo parleremo degli attacchi il cui scopo è quello di spiare una conversazione VoIP, mentre nel settimo capitolo illustreremo le tecniche utilizzate per intercettare e modificare il traffico di una conversazione. Nell'ottavo capitolo tratteremo dell'autenticazione delle parti coinvolte in una comunicazione, mentre nel nono illustreremo una delle tecniche maggiormente utilizzate per proteggere un'infrastruttura VoIP: le Virtual LAN. Infine, nel decimo e ultimo capitolo parleremo di come la crittografia possa essere utilizzata per garantire la segretezza di una conversazione.

# Capitolo 1

# Le reti di telecomunicazioni

## 1.1 Introduzione

Una rete di telecomunicazione è un complesso di mezzi implementante telecomunicazioni tra più di due persone situate in più di due diversi punti spaziali. In questo capitolo daremo una panoramica delle reti di telecomunicazioni, parlando prima delle reti di dati e successivamente di quella del sistema telefonico tradizionale.

## 1.2 Reti di dati

La teoria ed il conseguente sviluppo delle reti di telecomunicazioni è una delle realtà più complesse mai realizzate dall'uomo. Il networking affonda le sue radici nella struttura degli strumenti di comunicazione più disparati (in gergo denominati protocolli) per stabilire connessioni tra calcolatori o più in generale tra entità multi livello.

Realtà o filosofia astratta? Oggi, nell'era della Profit Economy, più che mai concretezza: le entità di cui parliamo possono essere applicazioni software che sono direttamente visibili all'utente finale, oppure programmi scritti per agire da intermediari (middleware) di cui spesso ignoriamo l'esistenza, ma che costituiscono quella architettura multi livello che permette di gestire servizi complessi.

Come il processo di realizzazione di un'automobile passa attraverso sottoprocessi più elementari che scompongono il problema nel suo insieme, la costruzione dell'autoveicolo appunto, così una rete di telecomunicazioni o la stessa Internet ed i suoi servizi sono l'incastro di elementi articolati che consentono il funzionamento dell'intero sistema.


Figura 1.1: Architettura dei protocolli

Figura 1.2: Una rete LAN di tipologia BUS

## 1.3 Reti locali

Il confine tra reti geografiche e locali non rappresenta una demarcazione caratterizzabile con qualche formula matematica o artificio linguistico. Possiamo pensare ad una rete locale come una realtà che consente il trasporto delle informazioni in aree circoscritte, come edifici di aziende o enti pubblici, campus universitari, fiere, centri turistici o agglomerati industriali.

Lo standard de facto sicuramente più utilizzato ad oggi, quando parliamo di reti locali, è costituito dai protocolli della suite IEEE 802, che ripercorrono la trasmissione dati su cavo via Ethernet (802.3) oppure tramite i sistemi wireless WiFi (802.11). Il protocollo Ethernet 802.3 fornisce uno standard per l'accesso al cosiddetto "mezzo condiviso", rappresentato fisicamente dai 4/8 fili dei cavi UTP (Unshielded Twisted Pair) (esiste anche l'analogo per la fibra).

Questo tipo di rete locale consente di attestare molteplici apparati

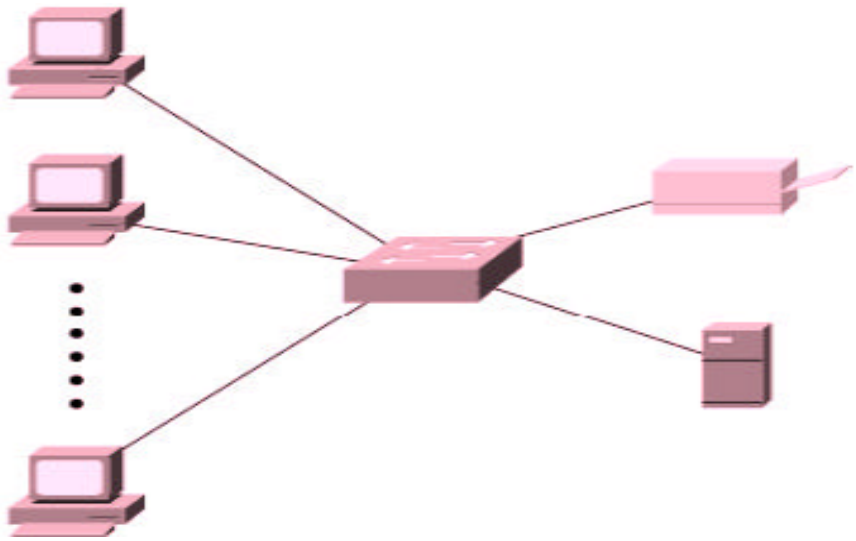

Figura 1.3: Realizzazione di una LAN con switch

(client ethernet) su un unico BUS: i metodi per realizzare tutto ciò sono in genere basati su switch, cioè su macchine in grado di strutturare le interconnessioni al BUS in appositi centri stella dove convergono i cavi UTP delle varie schede di rete.

La condivisione di un unico mezzo fisico da parte di molti utenti è resa possibile grazie ad un protocollo derivato dai sistemi CSMA/CD (Carrier Sense Multiple Access Collision Detect) in grado di gestire correttamente il problema della collisione dei pacchetti dati. Volendo semplificare, questi strumenti sono basati su una parte di Carrier Sense Multiple Access (CSMA), cioè sul controllo della portante o meglio del segnale per capire se qualcun altro sta già trasmettendo. Tuttavia questa prima fase da sola non elimina la possibilità che un client ethernet ascolti, non rilevi alcuna presenza di altri segnali e nel frattempo qualcun altro esegua le stesse operazioni.

In tal caso ci ritroveremo di fronte a due trasmissioni sullo stesso mezzo fisico e quindi ad una collisione di informazioni (o frame come vengono denominate in genere le unità informative ethernet). In pratica, nell'eventualità che la rete sia libera e due o più macchine inizino a trasmettere frame contemporaneamente, i segnali inviati da ciascuna stazione, interferendo vicendevolmente, distorcono i segnali originali.

Ecco che interviene la seconda parte del protocollo: il Collision De-

tection (CD). Per evitare queste collisioni, le schede di interfaccia, oltre a trasmettere, restano in ascolto tramite un circuito per controllare se durante la trasmissione si siano verificate delle interferenze, per poi ritrasmettere il frame non appena la rete risulti nuovamente libera. Le reti WiFi stanno attualmente trovando grosse prospettive di espansione in virtù degli indiscutibili vantaggi che offrono rispetto ai cablaggi tradizionali: minore necessità di punti rete fissi e di stesura cavi, ridottissimo impatto ambientale ed elettromagnetico, migliore gestione della mobilità e generale miglioramento nei processi di lavoro.

Anche le reti wireless sfruttano protocolli di accesso al mezzo condiviso come succede su ethernet ed implementano tutta una serie di accorgimenti per adattare al meglio il trasporto dell'informazione su etere: crittografia e controllo degli accessi per garantire sicurezza, controllo di errore, gestione dei segnali wireless multipli (ultimamente anche di quelli riflessi) per la miglior ricetrasmissione possibile. Parlare di protocolli Ethernet e Wireless soltanto in termini di Local Area Network (LAN) è diventato però riduttivo ad oggi. Sempre maggiore è infatti la disponibilità di sistemi che consentono di estendere l'utilizzo di tali tecnologie anche su tratte geografiche. Si pensi ad esempio alla possibilità attuale di implementare reti private virtuali (VPN) in grado di realizzare una vera e propria LAN geografica ed associarla a realtà aziendali multi sede che sentono sempre maggiore la necessità di interconnettere le varie località operative della stessa compagnia per trattare il trasferimento delle informazioni tra una sede e l'altra come se fossero virtualmente un'unica sede. Anche i sistemi WiFi e Wireless in genere stanno assumendo connotazioni che non sono più limitate a zone ristrette, ma costituiscono vere e proprie architetture nelle dorsali delle reti geografiche o di ultimo miglio.

## 1.4 La rete telefonica tradizionale

La rete telefonica tradizionale viene chiamata PSTN (Public Switched Telephony Network) ed è stata pensata e progettata per soddisfare le due esigenze principali della telefonia e cioè la trasmissione (il trasporto) della voce e la commutazione (il routing o instaurazione della chiamata). Quando un utente compone un numero telefonico ed entra in comunicazione con un altro utente, le centrali componenti la rete PSTN instaurano un circuito fisico temporaneo dedicato a quella comunicazione. Vengono

quindi allocate staticamente delle risorse di rete per costituire un circuito fisico che permetta la comunicazione, per questo motivo la rete PSTN viene anche chiamata "rete a commutazione di circuito". Al contrario le "reti a commutazione di pacchetto" instaurano dei circuiti virtuali senza dedicare in maniera esclusiva delle risorse ai due capi della comunicazione.

I vantaggi della PSTN sono quelli di garantire una buona qualità della "conversazione telefonica" che risulta così garantita e costante, di essere una rete capillare in grado di raggiungere la totalità della popolazione, di utilizzare una tecnologia stabile e consolidata e di poter disporre di personale con know-how adeguato. Per contro, esistono ad oggi delle mancanze e dei limiti oggettivi, che con l'avvento delle nuove tecnologie insidiano lo strapotere delle reti tradizionali. Le reti PSTN sono di fatto gestite da operatori dominanti, gli investimenti necessari a realizzare una rete capillare sono proibitivi, l'utilizzo della rete non è ottimizzato, l'allocazione statica dei circuiti è uno spreco (le statistiche parlano di un utilizzo inferiore al 50% della banda disponibile), infine, l'evoluzione tecnologica è orientata alla realizzazione di reti convergenti (voce, dati e video) che potranno essere realizzate utilizzando protocolli basati su IP e le tecnologie Voice Over Internet Protocol (VoIP).

# Capitolo 2

# Introduzione al VoIP

## 2.1 Introduzione

Il business nel gestire e rendere sicuri i nostri dati privati sta diventando molto importante ai nostri giorni. I benefici della comunicazione elettronica vanno incontro di pari passo, in proporzione, a rischi inerenti la sicurezza. I sistemi business critici possono essere compromessi regolarmente e usati per scopi illegali. Ci sono diversi esempi a riguardo: Seisint (ricerca Lexis-Nexis), Choicepoin, Bank of America, PayMaxx, DSW Shoe Warehouses, Ameriprise, and T-Mobile sono tutti esempi recenti (anno 2005).

Ad esempio Seisint nel 2005 (Lexis-Nexis research) è stato compromesso, sono stati compromessi nomi, indirizzi, sicurezza della società stessa, e informazioni su licenze di guida relative a 310.000 persone.

Una cosa che hanno in comune questi sistemi è che si basano sui dati di rete.

La pratica della sicurezza delle informazioni è diventata sempre più complessa da gestire, secondo la previsione di Gartner, una società su cinque ha una LAN senza fili e per il 60% delle WLAN non sono presenti nelle funzioni di base meccanismi di sicurezza attivi.

In questo capitolo forniremo una panoramica generale sulla tecnologia VoIP, illustrando quelli che sono i vantaggi che si traggono dal suo utilizzo.

## 2.2 Voce & Dati : due mondi separati?

La telefonia è la comunicazione di informazione vocale tra due o più partecipanti, mediante dei segnali trasportati sulle onde elettriche o ra-

dio waves. Fin da quando Alexander Graham Bell inventò il circuito telefonico e per primo prevedette il sistema telefonico pubblico, utenti e imprese hanno contato sul telefono come il pezzo forte per l'interazione umana. Con l'avvento delle tecnologie Internet e la connettività a banda larga, una nuova famiglia di tecnologie telefoniche ha iniziato il suo cammino: Voice over IP (VoIP). VoIP ha un grande significato per le aziende, per i service providers e per i clienti, perchè permette il trasporto di chiamate vocali, videoconferenze, e altre applicazioni real-time. VoIP può rimpiazzare il vecchio sistema telefonico, o può aggiungere valore allo stesso sistema telefonico tradizionale. Per esempio, la connettività a lunga distanza tra due uffici con il sistema telefonico tradizionale può spesso essere abbinato ad un basso costo per le chiamate quando è gestito con VoIP. Una dozzina di standard definiscono come lavora Voice over IP, ma esiste una piccola documentazione sulle migliori pratiche per implementare e gestire la tecnologia in azienda. Ci sono stati molti errori d'implementazione ad alto livello tra le grandi aziende che l'hanno adottato, per cui ci si spiega il motivo per il quale la telefonia IP ha una reputazione tanto indimidatoria. Tuttavia, se fatto bene, VoIP può trasformare il modello di costo delle telecomunicazioni combinando l'esperienza dell'overhead della voce e dati e le loro infrastrutture, può aumentare la produttività per gli utenti finali introducento nuove funzionalità e per gli amministratori telefonici gestendo in modo centralizzato le funzioni. VoIP può decrementare l'espansione dei progetti di integrazione della futura computer-telephony, mentre collega facilmente sistemi voce con webservers e database application.

## 2.3 Lo switch lascia le basi tradizionali

Le reti telefoniche sono state progettate per le trasmissioni vocali, le data networks no. Recentemente, negli ultimi anni, le funzionalità del PBX sono cambiate logicamente (e quindi anche fisicamente) dallo spazio chiuso allo spazio della rete esterna. La maggior parte delle persone considera la tradizionale rete telefonica pubblica commutata (PSTN) sicura. Sulla PSTN l'utente maligno, per attaccare il sistema deve accedere fisicamente alla linea telefonica e deve essere dotato di un adeguato hardware, di un dispositivo che genera bug. Sebbene sia possibile compromettere SS7, solo sofisticate tecniche e accessi diretti alla segnalazione del canale lo rendono possibile. Diversamente, i più importanti standard nel data networking, ad esempio TCP/IP(Transmission Control Proto-

col/Internet Protocol), sono relativamente stabili da più di 20 anni. C'è un alto grado di inconsistenza nel supporto e nell'implementazione degli standard per VoIP, dovuto per la gran parte all'evoluzione degli standard stessi e in parte ai vendor che tentano di bloccare le implementazioni non comuni delle implementazioni del protocollo. La conseguenza di questo è che in alcuni casi, applicazioni appena sviluppate, e quindi poco sicure, riempiono il mercato. Un'altra differenza sostanziale tra VoIP e gli altri protocolli è che i protocolli VoIP tendono a dividere in canali differenti la parte riguardante la segnalazione e dall'altra l'invio di data stream. Questi canali girano su combinazioni dinamiche Indirizzo IP/porta e questo causa una importante constatazione riguardante la sicurezza: se si combina questa separazione con la naturale realtà dove gli utenti si aspettano di essere semplicemente in grado di fare chiamate sia in entrata che in uscita, allora si dovrebbe cominciare a capire che progettare VoIP è più impegnativo e necessita di maggior sacrificio per garantirne tecnicamente la gestione delle chiamate.

## 2.4 Limiti della linea telefonica tradizionale

Le capacità del PSTN (Public Switched Telephone Network) sono esattamente proporzionali alla sua connessione fisica. Poichè ogni chiamata deve avere un circuito, un loop settato all'inizio della chiamata e chiuso alla fine, le attrezzature della PSTN svolgono molto lavoro, quindi ci sono delle grandi limitazione associate alla natura del cosiddetto "circuit-switched".

Potrebbero essere inserite nuove funzioni dalla compagnia telefonica, ma ci vorrebbero molti anni per fare l'upgrade delle centraline telefoniche per supportare queste funzioni, come ad esempio l'attesa di chiamata, o la chiamata a 3 utenti. Fino ad ora, alcune parti della PSTN non supportano l'ID del chiamante.

I limiti per le sue capacità sono un'altra sfida ingegneristica alla telefonia tradizionale. La bontà della riproduzione del suono di una chiamata è limitato alla banda disponibile tra il chiamante e il destinatario, e il massimo numero di chiamate tra due uffici è limitato alla disponibilità di circuiti voce che esistono tra di loro. Il problema posto all'azienda è in termini di costi, ed è il seguente: ogni circuito PSTN usato dall'azienda, può essere una linea POTS(Plain Old Telephone Service) o una T1, però

va aggiunta alle sue spese di telecomunicazione.

Le compagnie telefoniche e i vendor delle attrezzature telefoniche hanno fatto grandi passi in avanti per identificare e risolvere i problemi di costo e di capacità. Una grande densità di circuiti digitali come T1 e T3 hanno portato al ribasso il costo di un alta densità di telefonia, e le funzioni del PBX (Private Automatic Branch eXchange) come il Least Cost Routing (LCR) permettono all'azienda di minimizzare la sua grande spesa di chiamate a distanza.

Le chiamate a grande distanza sono diventate meno costose e il costo di un attrezzatura PBX caratterizzata da una grande attività di telefonia è scesa nel corso del tempo, di conseguenza.

Ad ogni modo, le funzioni telefoniche furono considerate un vantaggio competitivo. Quando le imprese le adottarono, esse diventarono parte del costo d'impresa, e gli utenti cercarono un nuovo paradigma di telefonia che potesse essere ispirato ad una maggiore competitività. La domanda a cui l'industria telefonica cerca di rispondere è : "dove possiamo arrivare partendo da quì?".

Gli innovatori delle aziende telefoniche guardarono ad Internet come risposta perchè il nucleo delle differenze nella filosofia ingegneristica e molti anni di discussioni accese sulla materia, portarono ad assumere che Internet è superiore in molti, troppi campi rispetto alla tradizionale rete telefonica.

Sulla rete Internet (e la rete IP in generale), i protocolli di comunicazione sono in uno stato costante di miglioramento, quindi possono essere scoperte sempre più nuove funzioni mentre l'efficienza di banda migliora costantemente e il costo della rete si abbassa.

Su Internet, la capacità è strettamente legata alla efficienza del software piuttosto che alla capacità fisica dello switch telefonico come il PBX e la PSTN. Così come un software migliora, la rete IP cresce nella sua capacità, mentre gli switch tradizionali hanno bisogno di hardware aggiuntivi per accrescere le loro capacità.

La rete IP ha ancora un altro vantaggio rispetto a quella tradizionale: il software usa componenti hardware standardizzati come PC a basso costo. Questo significa che fin quando gli upgrade dell'hardware sono necessari, possono essere procurati a basso costo. A differenza del tradizionale PBX, gli upgrade sulla rete IP intrinsecamente migliora la produttività e aumenta la capacità. Generalmente, la capacità è facilmente scalabile sulla rete IP piuttosto che sulla rete a commutazione di circuito come la PSTN. C'è da dire però che la PSTN è piuttosto affi-

dabile, è molto meno catastrofica rispetto alla rete IP mentre appunto l' Internet Protocol permette la ridondanza e le capacità di failover che sono poco costose e relativamente facili da implementare.

La differenza geografica, una tecnica usata sulla rete dati per eludere le interruzioni alla connettività locale, è molto facile per l'azienda ottenerla su Internet, ma è molto più difficile sulla PSTN. Per esempio potremmo connetterci a due Internet Service Provider e usare lo stesso insieme di indirizzi IP con entrambi, grazie al BGP standard, ma è quasi impossibile usare lo stesso insieme di numeri telefonici con due compagnie telefoniche.

Dato che la maggior parte delle reti delle aziende moderne usa gli stessi protocolli come Internet, c'è bisogno solo un arco di tempo prima che i vantaggi di questi protocolli diventino ideali per progettare le reti vocali. Il risultato di questa evoluzione è una immensa famiglia di tecnologia chiamata Voice Over IP, o VoIP.

VoIP è improriamente definita come una tecnologia che semplicemente usa la suite dei protocolli TCP/IP per facilitare la conversazione vocale, ma in realtà, è molto più di questo. Può essere usata per rimpiazzare la tradizionale linea telefonica nelle aziende o nelle case o anche per aggiungere nuove funzioni alla tradizionale linea telefonica. VoIP può inoltre risolvere i cambiamenti di connettività come collegare il PBX ai siti remoti, collegare estensioni della linea telefonica privata ad un singolo sito insieme al PBX, o semplicemente aggregare chiamate attraverso un piccolo telefono analogico come un key stream.

VoIP può essere usato per facilitare la comunicazione vocale anche in molti sottostrati applicativi. Può gestire le capacità delle chiamate on-demand agli utenti di una pagina web e permettere alla persone di usare il loro personal computer come se stessero usando le funzionalità della azienda telefonica.

## 2.5 Benefici del VoIP

Quali sono le promesse di convergenza per dati e voce sulla stessa infrastruttura fisica? In primo luogo, si può effettivamente ridurre i costi dopo tutto, grazie alla economia di sostenere una rete invece di due. VoIP, da un punto di vista gestionale, è meno costosa di quella di due infrastrutture di telecomunicazione. L'implementazione può essere costo-

sa e difficile, ma è ammortizzata con minori costi operativi e facilità di gestione e amministrazione. Il ritmo e le qualità di applicazioni basate su IP sta aumentando di pari passi con l'adozione del VoIP. Le caratteristiche che non erano disponibili a sistemi tradizionali, come ad esempio "click-to-talk" possono essere rapidamente modificate e distribuite. Ogni cifratura di voce che nel passato era limitata all'uso di selezionate compagnie, oggi può essere usata da tutti in un ambiente di sviluppo VoIP.

### 2.5.1 Ragioni per scegliere VoIP

1. Le periferiche VoIP sono facili da usare e poco costose nella manutenzione perchè fanno leva sulle reti di dati piuttosto che sulla rete che ha come solo scopo la voce.

2. VoIP aumenta il valore di Internet usando la comunicazione vocale.

3. Integrando la telefonia con le applicazioni multimediali, rende più semplice l'uso di VoIP rispetto ai sistemi tradizionali , perchè nelle impostazioni di VoIP, la gestione delle chiamate tende ad essere più aperta, standardizzata e software-driven.

4. VoIP può scalare molto più economicamente rispetto al sistema telefonico tradizionale perchè è spesso aggiornato con nuovi PC hardware.

5. Permette più controllo amministrativo centralizzato rispetto al tradizionale PBX.

6. La gestione di una rete VoIP è compiuta usando la stessa rete che prende le informazioni sulla voce, a differenza della PSTN che usa il protocollo SS7.

7. La resistenza ai guasti e ai periodi di interruzione della rete è più facile ottenerla con TCP/IP piuttosto che con il tradizionale sistema vocale.

8. Molte delle nostre attrezzature telefoniche esistenti possono interfacciarsi con i sistemi VoIP usando terminali analogici adattati, o ATA (Analog Telephone Adapter).

## 2.6 Protocolli VoIP

Due importanti suite multimediali ad oggi dominano il VoIP: SIP e H.323. Per semplicità, tratteremo SIP (Session Initiation Protocol) e H.323 (standard approvato dall'Unione Internazionale delle Telecomunicazioni - ITU - nel 1966 per rendere compatibili le trasmissioni in videoconferenza su reti IP) come protocolli di segnalazione.

Tuttavia, consideriamo che H.323 definisce esplicitamente i bassi livelli di segnalazione, mentre SIP è più un application-layer control framework. Le Request line di SIP e i campi header, definiscono il carattere della chiamata in termini di servizi, indirizzi, e caratteristiche del protocollo.

Il trasporto vocale, è trattato maggiormente da RTP (Real-time Transport Protocol) e RTCP (Realtime Transport Control Protocol), inoltre anche SCTP (Stream Control Transmission Protocol) è stato proposto e ratificato dall'IETF ed è usato per la versione SS7 (Signaling System 7) di IP, conosciuta come SIGTRAN (segnalazione dei trasporti).

Il trasporto della voce su IP richiede inoltre un gran numero di protocolli di supporto che sono usati per garantire QoS, provvedere al name resolution, permettere upgrade di firmware e software, sincronizzare i clock nella rete, fare efficientemente il routing delle chiamate, monitorare le performance e permettere l'attraversamento dei firewalls.

SIP è un protocollo di segnalazione per le conferenze Internet, telefoniche, notifiche di eventi, e instant messaging; è un IETF-ratified response-request protocol, dove il flusso dei messaggi ricorda l'HTTP (HyperText Transport Protocol); è un framework in cui il solo scopo è stabilire la connessione, non si interessa sui dettagli della chiamata, inoltre i messaggi SIP sono codificati con ASCII (American Standard Code for Information Exchange).

Dall'altra parte, H.323 è un ITU (International Telecommunications Union) protocol suite, più vicino alla filosofia del SS7. Lo standard H.323 provvede a fornire comunicazione audio, video e dati, su una rete basata su IP. I protocolli che formano l'H.323 sono compilati usando ASN.1 (Abstract Syntax Notation), PER (Packed Encoding Rules)—un sottoinsieme di BER (Basic Encoding Rules)— è un binario codificato, usato su reti con banda limitata. A differenza di SIP, H.323 definisce esplicitamente quasi ogni aspetto di una chiamata.

Entrambe le suite contano su protocolli supplementari al fine di fornire servizi ausiliari.

Utilizzano entrambi i protocolli TCP e UDP (User Datagram Protocol), ed entrambi aprono un minimo di cinque porte per ogni sessione di VoIP (Chiamata di segnalazione, due RTP e due RTCP). Entrambi i protocolli offrono caratteristiche comparabili, ma non sono direttamente interoperabili. Le portanti tendono a preferire H.323 perché i metodi definiti da H.323 fanno una traduzione dall' ISDN (Integrated Services Digital Network) o dal SS7 al VoIP più semplice che per SIP.

SIP, d'altra parte, è text-based, funziona meglio con i messaggi istantanei, e di solito è implementato su hardware meno costoso. Ad oggi, H.323 è stato il leader del mercato, ma SIP lo sta rapidamente soppiantando.

## 2.7 VoIP non è solo un altro protocollo dati

La telefonia IP usa l'architettura di internet, così come ogni tipo di applicazione dati. Tuttavia, in particolar modo da un punto di vista di amministratore di sicurezza, VoIP è leggermente differente in quanto ci sono anche altri aspetti da considerare. Ci sono tre motivi principali per questa affermazione:

- Le conversazioni vocali possono essere iniziate all'esterno del firewall. Molti client driven protocols iniziano la richiesta dall'interno del firewall. La figura 2.1 mostra il flusso base di un tipico scambio di messaggi.

- Data la natura real-time di VoIP, se il pacchetto arriva un secondo troppo tardi, è inutile.

- La separazione di data e signaling, le sessioni, in particolar modo le sessioni in entrata che definiscono informazioni di indirizzamento per i dati canale, non interagiscono bene con il NAT (Network Address Translation) e la cifratura stessa del segnale.

Nella figura 2.1, una richiesta è iniziata da un client sulla parte interna di un firewall su un demone server residente su un host esterno al firewall. I firewall che sono capaci di inspezioni di stato monitoreranno la connessione ed apriranno le porte interne (se la porta è associata con una sessione già stabilita). L'Application Layer Gateways (ALGs) potrà comportarsi in un modo simile, proxando le connessioni in entrata e in uscita per le richieste degli host interni. Per l'amministratore del firewall

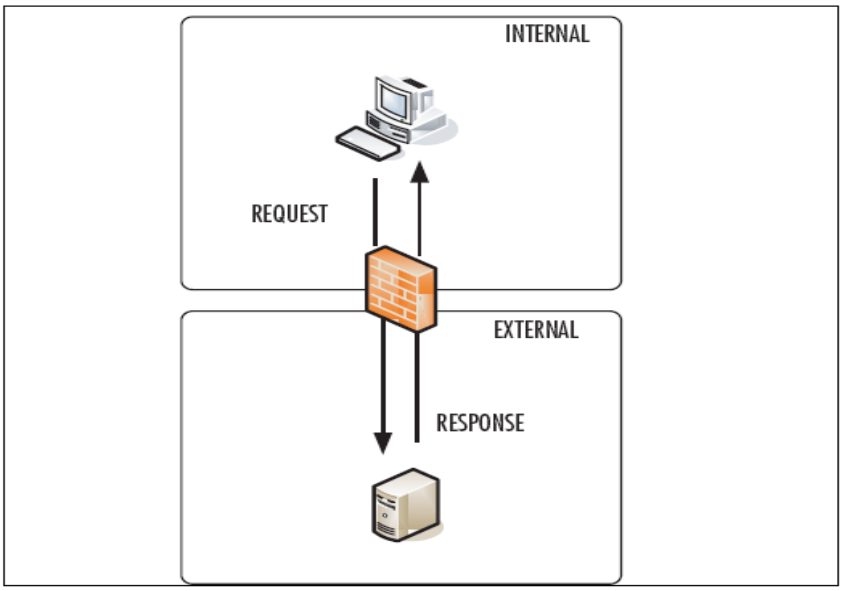


Figura 2.1: Normale flusso del messaggio

e per l'user, la sessione è completamente normale, ed è la massima sicurezza consentita dai firewall.

Nella figura 2.2, la topologia della request-response è diversa.

Ci sono altre differenze. La sensitività di VoIP alle condizioni negative della rete è molto differente quantitativamente dal tipo della qualità del traffico data. Le applicazioni Real-time, incluso VoIP, hanno dei requisiti molto più stringenti di un normale protocollo di trasporto come IP. Ogni pacchetto VoIP è formato da circa 20 ms di voce in media. Una singola perdita di pacchetto potrebbe non essere importante, ma la perdita di pacchetti multipli è interpretata dall'user come una qualità negativa. Gli ingegneri di rete sono abituati a brevi periodi di interruzione, ma la maggior parte degli utenti non prende di buon occhio le interruzioni, anche se le tollerano. Anche se i telefoni cellulari hanno la straordinaria caratteristica di maggiori connessioni cadute, nell'uso di IP gli utenti si aspettano che i loro telefoni lavorino per tutto il tempo senza interruzioni. La disponibilità è un elemento chiave per le metriche di prestazione VoIP.

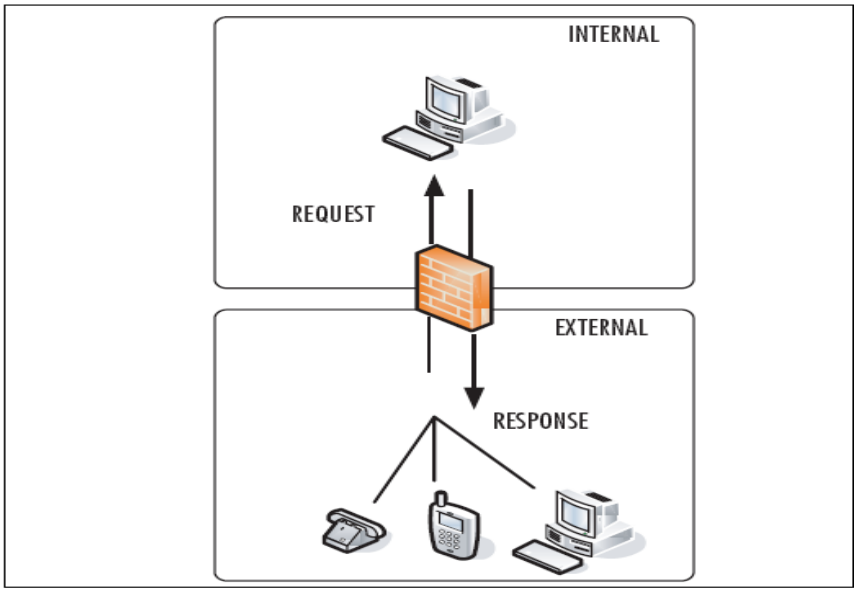


Figura 2.2: Flusso VoIP per i messaggi in entrata

# Capitolo 3

# La telefonia tradizionale e IP

## 3.1 Introduzione

Anche dopo l'introduzione del VoIP, le imprese di telefonia sono rimaste concentrate su due settori: (1) ridurre il costo del Public Switched Telephone Network (PSTN) e la connettività globale, (2) l'aggiunta di funzionalità di comunicazione. Dal momento che il PBX è stato introdotto nel 1879, i clienti hanno cercato risparmi sui costi riducendo il numero di linee fisiche con la rete PSTN.

Il lavoro della PSTN (Public Switched Telephone Network) è di facilitare le conversazioni telefoniche ad ogni ora. La PSTN combina collegamenti dati analogici, digitali e elettromagnetici per rendere sicura ogni giorno la possibilità di ricevere telefonate a qualsiasi orario.

Inizieremo il capitolo trattando dell'infrastruttura di rete del sistema telefonico tradizionale. Successivamente discuteremo i pricipali protocolli in uso oggi per un infrastruttura di telefonia basata su IP.

### 3.1.1 Mesh vs Switched

Quando progettiamo una rete che connette molti telefoni (endpoint) ci sono due approcci: le mesh networks e le switched networks. In una progettazione di una mesh networks, ogni endpoint ha una connessione permamente a qualsiasi altro endpoint, così tutti possono comunicare con tutti, le reti mesh non sono così pratiche però, perchè una volta che aggiungiamo dei nuovi endpoints, il numero dei collegamenti aumenta a dismisura fino ad essere una quantità assurda. Per esempio, in una rete mesh con 10 endpoints, devono essere gestiti 100 links, quindi, si preferiscono usare le switch network. Nelle switched network, i link tra gli endpoint non sono permamenti, infatti l'unico momento nel quale c'è

bisogno del collegamento tra i diversi endpoints è quando arriva una chiamata, per il resto del tempo il link è libero, non usato. Lo switching è un metodo nel quale i link sono stabiliti e rimossi quando c'è n'è bisogno, eliminando il bisogno della rete mesh. La PSTN è una rete switched.

La PSTN trasporta ogni chiamata telefonica settando e chiudendo un link temporaneo, spesso un circuito elettrico tra chiamato e chiamante. I links che prendono le chiamate possono essere compresi tra fili di rame, fibre ottiche o sistemi radio, dipende dal tipo di infrastruttura di rete che esiste tra chiamante e chiamato.
Mentre le connessioni telefoniche usate nelle maggior parte delle abitazioni sono tradizionalmente analogiche, elettromagnetiche, quelle tra grandi switches telefonici sono digitali.

### 3.1.2 Signaling System 7

SS7 è il componente di segnalazione della PSTN, una seconda rete che gira accanto alla PSTN nella quale lo scopo principale è coordinare la comunicazione tra gli switch, database delle compagnie telefoniche e sistemi di pagamento/fatturazione, e altre parti della rete vocale pubblica. Le chiamate dei numeri verdi e il routing delle chiamate a lunga distanza sono entrambe funzioni della SS7. Mentre la SS7 è secondaria alla principale funzione della PSTN (settare e chiudere i link sulla switched network), le moderne PSTN potrebbero non funzionare correttamente senza di essa.

Oggi, PSTN è il sistema di intercomunicazione più usato nel mondo. Per quanto riguarda la gestione del traffico vocale, non ha rivali. I servizi VoIP, come skype hanno puntato proprio su questo, il loro modello business fa forza sul rafforzamento delle interconnessioni PSTN. Il PSTN fornisce servizi FAX, data, telex, video e centinaia di altri servizi multimediali e per molti decenni, PSTN ha goduto di uno schema di numerazione universale chiamato E.164. In pratica, ogni volta che vediamo un numero che inizia per '+', stiamo vedendo un numero in formato E.164. Anche Internet stessa dipende dal trasporto PSTN per stabilire circuiti e chiamate.

Per poter ben definire PSTN oggi, dobbiamo concentrare l'attenzione su diversi aspetti. Primo aspetto, il livello fisico, l'impianto via cavo richiesto per la distribuzione del segnale, dal twisted-pair copper al coaxial electric alla fibra ottica. Secondo, i modelli di trasmissione del segnale stesso, modulazioni analogiche e digitali e trasmissioni su interfacce elet-

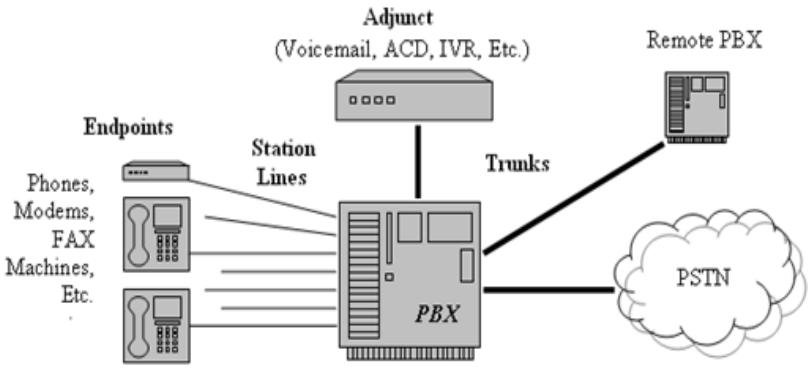


Figura 3.1: Un diagramma base di PBX

triche, ottiche, e radio. Terzo, i sempre più sofisticati meccanismi di protocolli di segnalazione (di controllo), la cosiddetta "rete intelligente" introdotta con gli Integrated Service Digital Network (ISDN).

### 3.1.3 Sistemi PBX tradizionali

Quando il telefono è diventato molto importante nel business mondiale, gli innovatori hanno esteso le sue capacità e lo hanno reso più conveniente. Essi l'hanno fatto usando periferiche gateway per telefonia che connettono privatamente tra di loro i proprietari di telefoni in una rete vocale privata con funzioni di chiamata auto-organizzate. In molti grandi uffici, i telefoni si connettono ad un privato, con lo switch che interfaccia le linee della compagnia telefonica, questo switch è chiamato PBX. Uno dei compiti del PBX è determinare come fare il routing delle chiamate, cioè, come verificare quando la chiamata sta provando a raggiungere un' altra persona nello stesso ufficio, o sta provando a raggiungere qualcuno via PSTN.

Le compagnie telefoniche hanno usato il private branch exchange (PBX) per più di un secolo, quindi è facile capire il perchè il VoIP venga considerato una alternativa moderna al tradizionale PBX. Tuttavia, questa confronto è sbagliato, perchè il centralino PBX come concetto è neutrale a livello di trasporto. Nella figura **??** viene mostrato un normale centralino telefonico

Vediamo ora una panoramica tra i due principali protocolli dedicati al VoIP e cioè H.323 e SIP per comprendere quali differenze producono e con quale ottica sono di supporto al VoIP.

## 3.2 H.323

L'H.323 è una raccomandazione ITU che fornisce le specifiche per definire un'infrastruttura per la trasmissione di dati multimediali, quali audio e video, su una rete a commutazione di pacchetto quale ad esempio Internet. H.323 fa parte di un numero più ampio di standard (H.32x) che si occupano della comunicazione su diversi tipi di rete, in particolare:

- H.310 - per comunicazioni multimediali su BISDN (Broadband Integrated Services Digital Network).
- H.320 - per comunicazioni multimediali su ISDN (Acronimo di Integrated Services Digital Network) a banda stretta.
- H.321 - per comunicazioni multimediali su ATM (Asynchronous Transfer Model).
- H.322 per comunicazioni multimediali su LAN.
- H.324 per comunicazioni multimediali su PSTN (Public Switched Telephone Network).

### 3.2.1 Architettura

La struttura di un generico sistema H.323 si basa sulle seguenti componenti:

- Terminali (Tx).
- Gateway (GW).
- Gatekeeper (GK).

**Terminali** Un Terminale H.323 è un endpoint interconnesso sulla rete in grado di comunicare con un altro Terminale o con un Gateway. Il contenuto della comunicazione può essere audio, video, filmati a colori e dati. Un terminale deve essere in grado di codificare e decodificare sia il parlato, secondo le specifiche ITU e CCITT (G.722, G.728, G.729, MPEG 1 audio e G.723.1), che il video (H261 e QCIF). Come componente H.323 deve poter gestire la trasmissione dei messaggi H225 e H245.

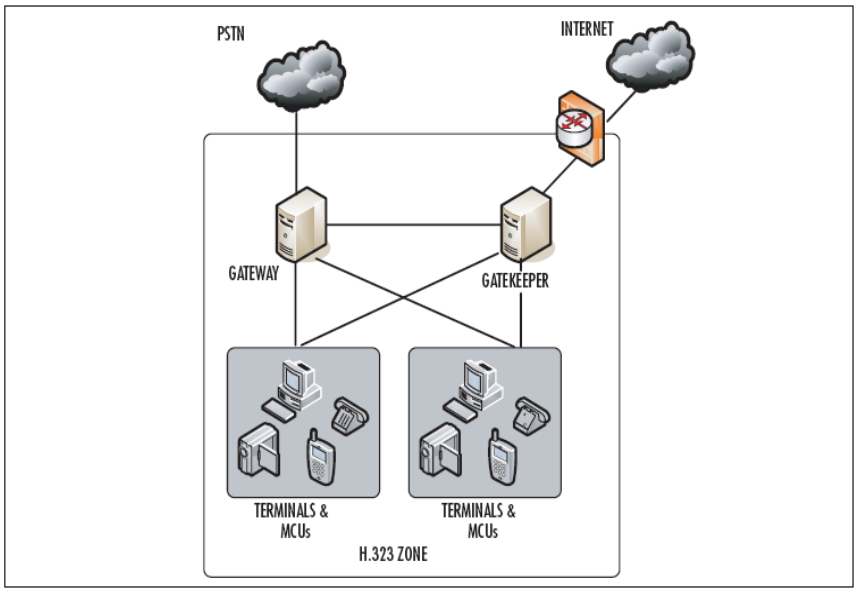


Figura 3.2: Entità H.323

**Gateway** Un voice gateway, anche definito come media gateway, è un apparato posto sulla rete che fornisce comunicazione bidirezionale e real time tra un terminale H.323 sulla LAN e un altro terminale H.323 posto anch'esso su un altro segmento di rete. Su un gateway possono essere presenti più interfacce per connettere diversi dispositivi come telefoni analogici o digitali (ISDN), in tali casi è il gateway a svolgere la funzione di terminale per questi telefoni.

**Gatekeeper** Il gatekeeper viene definito come un'entità H.323 sulla rete che si occupa di controllare lo stato delle chiamate in corso e l'utilizzo della banda. Ogni gatekeeper si occupa di una zona in cui sono presenti degli endpoint e cioè telefoni, terminali H.323 e voice gateway.

### 3.2.2 I principali protocolli connessi ad H.323

H.323 è specificato come un ombrello che comprende un grande numero di stati-macchina che interagiscono in diversi modi, dipendendo al momento stesso dalla presenza, assenza, e relazione topologica delle entità partecipanti e del tipo di sessione (per esempio audio o video). Ci sono molti sottoprotocolli nella specifica H.323. Per capire complessi-

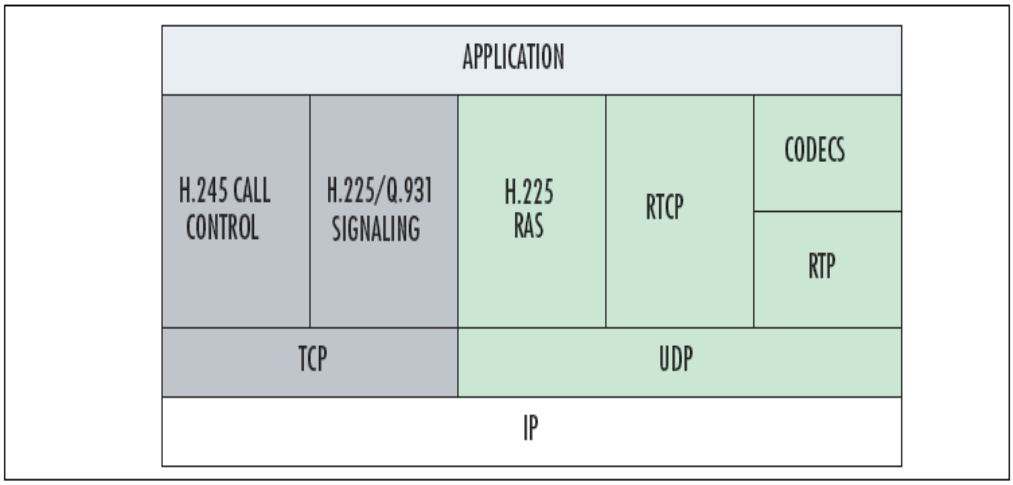


Figura 3.3: Stack dei protocolli collegati ad H.323

vamente il flusso messaggi in una transazione del protocollo H.323 ci si baserà sui protocolli più rilevanti nell'uso specifico di VoIP, nella figura **??** vengono mostrate le relazioni più importanti tra i maggiori protocolli VoIP sotto H.323.

H.323 definisce un insieme generale di procedure per il set-up e per la negoziazione della chiamata. I protocolli più importanti sono H.225, H.235, H.245, e i componenti della serie di segnalazione del protocollo Q.900. I metodi di trasporto base sono definiti dai protocolli real time RTP e RTCP. H.323 specifica inoltre un gruppo di codec audio per la comunicazione VoIP, la serie G.700:

- H.225/Q.931. Definisce la segnalazione per il setup e la chiusura della chiamata, includendo indirizzo IP di sorgente e destinazione, porte, country code, e informazioni sulla porta H.245.
- H.225.0/RAS. Specifica i messaggi che descrivono la segnalazione, ammissione della registrazione e status (RAS) e informazioni media stream.
- H.245. Specifica i messaggi che negoziano le capacità dell'insieme dei terminali, le relazioni master/slave, e informazioni sui canali logici per il media stream.
- Real Time Protocol (RTP). Descrive il trasporto end-to-end dei dati real-time.

- Real Time Control Protocol (RTCP). Descrive il monitoraggio della consegna dati e QoS fornendo informazioni come distorsione del segnale e media pacchetti persi.

- Codecs. La serie G.700 dei codecs usati per VoIP includono:

  – G.711. Uno dei più vecchi codec, G.711, non usa compressione, quindi la qualità della voce è eccellente, d'altra parte però, questo codec è quello che consuma più banda (è lo stesso codec usato da PSTN e ISDN).

  – G.723.1. Questo codec è stato progettato per la telefonia/videoconferenza sui linee telefoniche standard ed è ottimizzato per la codifica/decodifica veloce, ha una media qualità della voce.

  – G.729. Questo codec è usato principalmente nelle applicazioni VoIP perchè è quello che richiede minor uso di banda.

La segnalazione di H.323 tipicamente scambiata è inoltrata via gatekeeper o direttamente tra i partecipanti a seconda della scelta del gatekeeper, la comunicazione di tipo media invece è normalmente scambiata inoltrandola direttamente tra i partecipanti di una chiamata. La comunicazione dati H.323 usa entrambi i protocolli TCP e UDP. TCP assicura la comunicazione affidabile per segnali di controllo e dati, perchè il segnale deve essere ricevuto in ordine e non deve andare perso, UDP invece è usato per il segnale audio e video che è time-sensitive ma non sensibile alla perdita di pochi pacchetti, di conseguenza, il canale di segnalazione della chiamata H.225 e il controllo di chiamata H.245 non possono girare su TCP, mentre audio, video, e RAS sono affidati a UDP.

### 3.2.3 H.225/Q.931. Segnalazione di chiamata

Assumendo una procedura di "partenza lenta", il protocollo H.225 definisce due importanti fasi per il setup della chiamata: segnalazione di chiamata e RAS. La segnalazione di chiamata descrive gli standard per il setup della chiamata, manutenzione e controllo e chiusura. Un sottoinsieme dei messaggi di segnalazione di Q.931 sono usati per iniziare la connessione tra i due endpoint H.323 sui quali possono essere trasportati dati real-time. Il canale di segnalazione è aperto tra un endpoint-gateway, un gateway-gateway, oppure gateway-gatekeeper ancor prima di stabilire

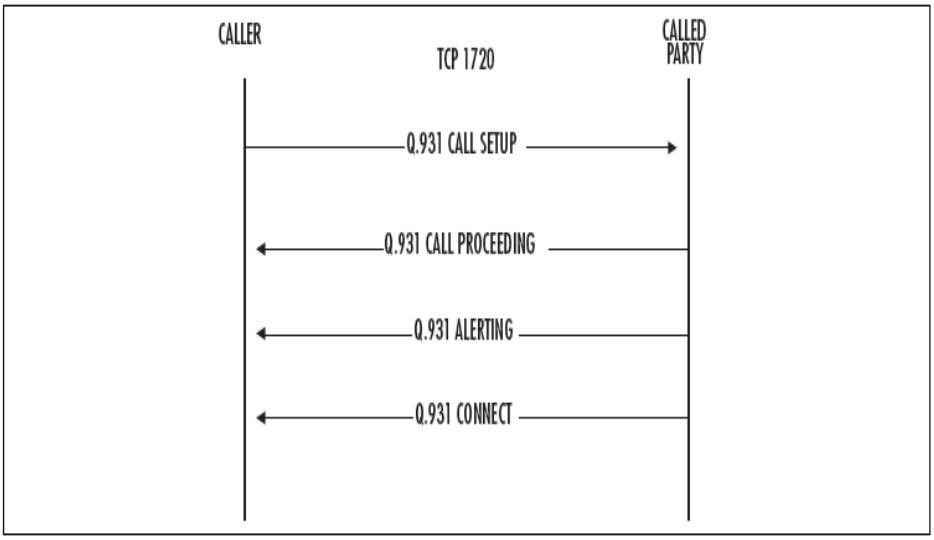


Figura 3.4: Segnalazione H.225/Q931

un qualsiasi altro canale; se non ci sono gateway o gatekeeper presenti, i messaggi H.225 sono scambiati direttamente tra i due endpoint. Il protocollo H.225 definisce inoltre i messaggi usati per la comunicazione tra endpoint-gatekeeper e gatekeeper-gatekeeper, e questa parte del protocollo H.225 è chiamata RAS (Registration, Admission, Status) e diversamente dalla segnalazione, gira sul protocollo UDP.

RAS è usato per compiere la registrazione, il controllo di ammissione, il cambiamento dello status di banda, e le precedure di chiusura tra gli endpoint e i gatekeeper. Lo stabilimento di una chiamata tra due endpoint, richiede un diverso programma di connessioni che dipende da quali entità sono coinvolte nella sessione. Per le connessioni dirette tra endpoint, vengono settati due canali TCP, uno per il setup della chiamata (Q.931/H.225) e un altro per la scambio di capacità e controllo di chiamata (H.245). La cosa che fa un endpoint è iniziare uno scambio H.225/Q931 su un canale TCP sulla porta ben conosciuta (TCP 1720) con un altro endpoint.

Se un gatekeeper è presente tra i due endpoint allora la segnalazione H.225 RAS precede la segnalazione Q.931 così come mostrato nella figura **??**

Questi messaggi sono usati per registrarsi ad un gatekeeper e per richiedere il permesso di iniziare la chiamata:

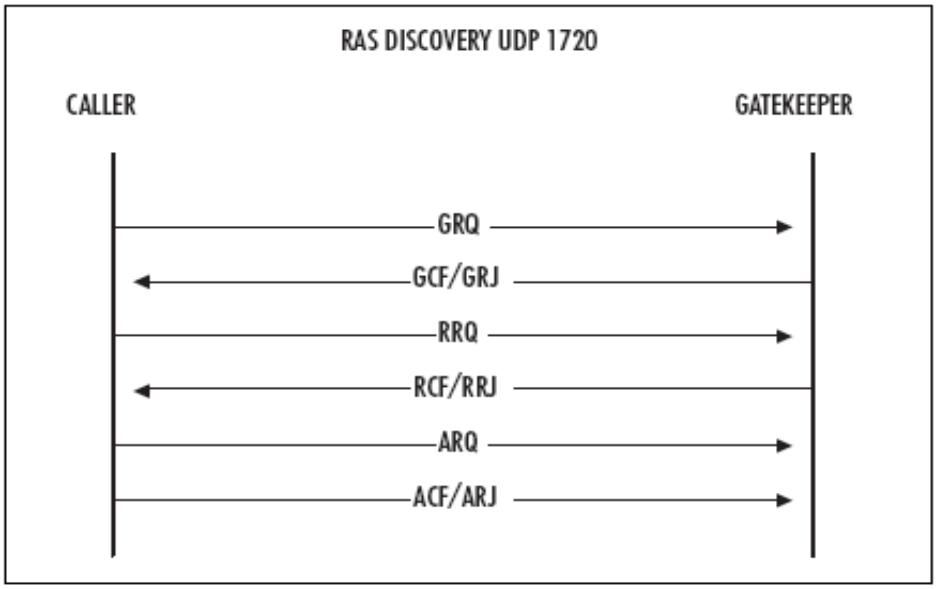


Figura 3.5: H.225/Q931 RAS

- Gatekeeper Request (GRQ). Il pacchetto GRQ è mandato unilateralmente per trovare qualche gatekeeper esistente. Questo richiede che gli indirizzi ip dei gatekeeper siano configurati sull'endpoint, se non lo sono, l'endpoint fa una ricerca multicast per trovare il gatekeeper.
- Gatekeeper Confirm or Reject (GCF/GRJ) Reply La risposta dal gatekeeper all'endpoint che rifiuta/conferma la richiesta di registrazione dell'endpoint.
- Registration Request (RRQ) Request Richiesta di registrazione da parte di un terminale o gateway ad un gatekeeper.
- Registration Confirm or Reject (RCF/RRJ) Il Gatekeeper conferma o rifiuta la registrazione.
- Admission Request (ARQ) Serve per richiedere l'accesso ai pacchetti della rete da un terminale al gatekeeper.
- Admission Confirm or Reject (ACF/ARJ) Il Gatekeeper può confermare o rifiutare l'ammissione; se conferma, l'indirizzo di trasporto e la porta usata per la segnalazione di chiamata sono inclusi nella risposta.

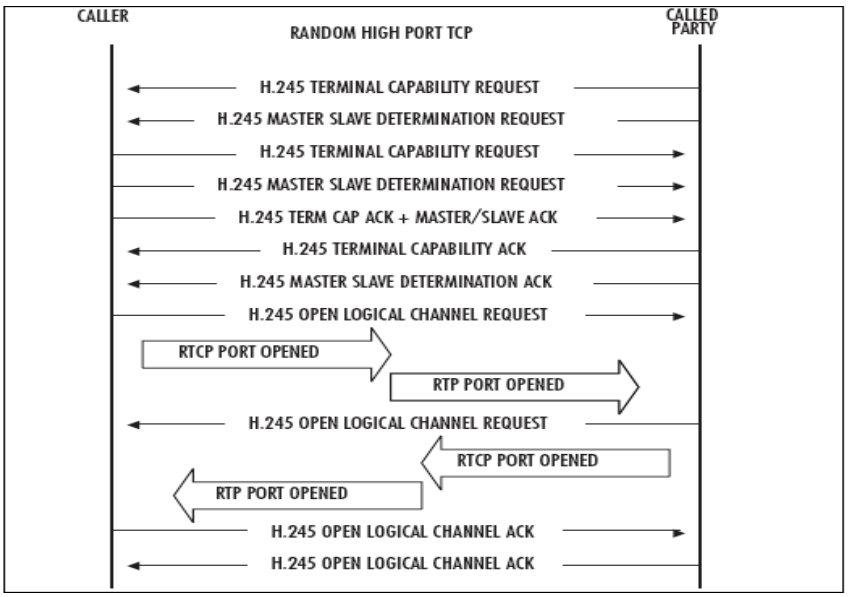


Figura 3.6: H.245 - Controllo della chiamata

### 3.2.4 H.245 Messaggi di controllo della chiamata

Dopo che una connessione è stata settata dalla procedura di segnalazione di chiamata, i messaggi H.245 sono usati per il tipo di chiamata multimediale, per scambiare le capacità dei terminali e per stabilire il flusso di chiamata prima che la stessa sia stabilita. H.245 gestisce inoltre i parametri della chiamata dopo averla stabilita. I messaggi H.245 inviati includono notifiche delle capacità dei terminali e i comandi per aprire e chiudere i canali logici. La negoziazione di H.245 prende luogo su canali separati l'uno dall'altro, dove in uno c'è lo scambio H.225 ma ultimamente le applicazioni supportano il tunneling di H.245 nel canale di segnalazione H.225. L'indirizzo di trasporto di H.245 è sempre passato nel messaggio di segnalazione della chiamata, cioè, le informazioni sulla porta sono passate nel payload che precede lo scambio H.225/Q.931. La figura 3.6 mostra un esempio di chiamata di controllo di H.245

Nel seguito della chiamata si aprono le porte TCP per stabilire il controllo del canale dopo aver estratto le informazioni dal H.225/Q.931 sul pacchetto di segnalazione. Durante questo scambio, le capacità dei terminali vengono negoziate, ad esempio i codec scelti e la determinazione del master/slave. La negoziazione multimediale inizia con il pacchetto OpenLogicChannel, quando si è pronti per parlare, c'è una risposta

OpenLogicChannel Ack che contiene le informazioni sulle porte dinamiche nel payload. Di fianco a questo, l'uso di porte dinamiche rende difficile l'implementazione di politiche di sicurezza su firewall e NAT. In alcuni casi dei particolari firewall H.323-aware o componenti firewall chiamati Application Layer Gateway (ALG) vengono relegati a far passare la segnalazione di H.323 e i media associati. Qundi entrambi i canali RTP/RTCP sono aperti e la comunicazione avviene.

## 3.3 Il protocollo SIP

Il protocollo SIP è nato per fornire un'alternativa ad H.323 nella gestione del Voice Over IP, vista la complessità di quest'ultimo. H.323 rappresenta infatti una serie di raccomandazioni sulla modalità di gestione del Voice Over IP, ma le specifiche che detta non sono facili da seguire. H.323 sfruttando protocolli esistenti, invece che specificarne di nuovi, rappresenta un sistema chiuso che non si integra facilmente con altre applicazioni. Si è così rivelato valido per le aziende che avevano appoggiato la sua creazione ma inadatto in molti altri casi.
IETF (l'Internet Engineering Task Force) ha quindi deciso di specificare un altro protocollo, più semplice e flessibile. Il risultato è stato il protocollo SIP che è rappresentato da un canale di comunicazione per lo scambio di brevi messaggi in formato testo in grado di gestire la segnalazione della chiamata ed il controllo di questa. Il protocollo non specifica invece come devono essere realizzati eventuali gateway per la connessione di dispositivi SIP ad altre apparecchiature.
A differenza di H.323 il numero di porte per le varie segnalazioni viene ridotto eliminando parte dell'overhead relativo agli header IP e il tempo necessario per instaurare la connessione. Il fatto che le segnalazioni avvengano tramite lo scambio di messaggi di testo rende possibili implementazioni del protocollo in maniera semplice ed aiuta il debug da parte degli sviluppatori.
Se H.323 fornisce delle direttive precise fin dalla sua nascita, SIP si è adeguato continuamente alle modifiche che si sono rese necessarie e si presenta quindi oggi come un sistema in evoluzione. Questo rappresenta, d'altro canto, anche il principale svantaggio di SIP e cioè la difficoltà di interoperabilità tra client che supportano diverse versioni del protocollo.

### 3.3.1 Utilizzi di SIP

Il formato dei messaggi scambiati e dei numeri SIP è stato influenzato da altri protocolli Internet. Un client SIP è infatti identificabile da una URL di questo tipo sip:user@domain.com, in questo modo è possibile effettuare una chiamata da una pagina web semplicemente registrando nel browser una applicazione in grado di gestire il protocollo SIP e "cliccando" sul link.
Lo stesso scambio di messaggi ascii lo accomuna ad HTTP o SMTP, mentre la formattazione del testo a richieste HTTP. I servizi offerti dal SIP sono:

- Trasferimento di chiamata nei casi di occupato, mancata risposta, senza condizioni, manipolazioni di indirizzo (come nelle chiamate di tipo 700, 800 , 900).
- Informazioni sul chiamante ed il chiamato.
- Mobilità del client, tramite registrazioni successive o redirezione della chiamata verso un proxy.
- Autenticazione del chiamante e del chiamato.
- Invito per comunicazioni multicast.
- Distribuzione base delle chiamate (ACD).

Tra le applicazioni che SIP può gestire ci sono:

- conferenze multimediali su Internet.
- chiamate telefoniche su Internet.
- distribuzione di contenuti multimediali.
- registrazione a servizi.
- monitoraggio.

Per svolgere questi servizi un client SIP deve essere in grado di creare, modificare e terminare sessioni con uno o più partecipanti. Queste sessioni possono gestire una qualsiasi delle applicazioni controllabili attraverso il protocollo SIP.

### 3.3.2 Gestione comunicazione tra User Agent

I partecipanti in una sessione possono comunicare utilizzando multicast (supportato solo marginalmente sulla rete Internet e simulato utilizzando mbone) oppure più relazioni unicast. Gli inviti SIP utilizzati per creare sessioni SIP trasportano anche le descrizioni delle sessioni con informazioni che permettono ai partecipanti di convenire su un insieme di formati multimediali. SIP supporta la mobilità degli utenti attraverso proxying e la redirezione delle richieste alla locazione dell'utente. Gli utenti possono inoltre registrare la loro locazione corrente ad intervalli regolari rendendosi in questo modo sempre disponibili. SIP è stato creato in modo da essere indipendente dal livello di trasporto sottostante e supporta cinque tipi di comunicazioni per stabilire e terminare le comunicazioni multimediali:

- User location.
- User capabilities.
- User availability.
- Call setup.
- Call handling.

Esse esprimono nell'ordine la locazione attuale dell'utente, la disponibilità ad esempio in termini di codec sonori, ad essere chiamato, il setup della chiamata e la sua successiva gestione. Il protocollo SIP funziona nel seguente modo:

I chiamanti ed i chiamati sono identificati da un indirizzo SIP. Quando viene fatta una chiamata, il chiamante prima localizza il server appropriato e quindi invia ad esso una richiesta. L'operazione più comune è l'invito utilizzato per iniziare una chiamata, tali richieste possono sia raggiungere direttamente il destinatario, oppure, attraverso una catena di richieste inviate di proxy in proxy essere portate al destinatario.

### 3.3.3 Struttura dell'header del protocollo

Il protocollo è composto da una linea iniziale che specifica la release del protocollo SIP utilizzato, l'intestazione del messaggio, una linea vuota e il corpo opzionale del messaggio.

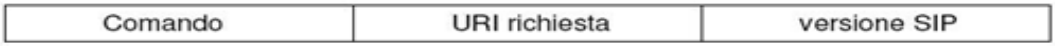


Figura 3.7: Intestazione del pacchetto SIP di richiesta

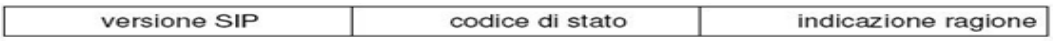


Figura 3.8: Intestazione del pacchetto SIP di risposta

Il formato dell'intestazione di una richiesta è rappresentato da:

**Comando**: tipo di messaggio che si intende inviare alla risorsa rappresentata dalla URI.

possibili comandi sono Invite, Ack, Options, Bye, Cancel, Register.

**Invite** Avvia una chiamata.

**Ack** Conferma di un messaggio ricevuto.

**Bye** Termina o trasferisce la chiamata.

**Cancel** Cancella ricerche o lo stato di ringing.

**Options** Caratteristiche supportate dall'inviante.

**Register** Registra presso un server.

Una SIP Uniform Resource Location (URL) oppure più genericamente una Uniform Resource Identifier (URI) rappresenta l'utente o il servizio a cui la richiesta viene inviata. Il formato di un messaggio di risposta è mostrato qui di seguito:

- Versione SIP: la versione SIP utilizzata.
- Codice di stato: il risultato del tentativo di interpretare e soddisfare la richiesta composto da 3 cifre decimali.
  - 1xx in ricerca, squillo, accodato.
  - 2xx successo.

  - 3xx forwarding.
  - 4xx errori lato client.
  - 5xx errori lato server.
  - 6xx occupato, rifiutato, non disponibile.

- Indicazione ragione: una descrizione testuale del codice di stato.

### 3.3.4 Entità protocollo SIP

Anche per il protocollo SIP sono state create delle entità con ruoli ben precisi per organizzare e strutturare insiemi di client.
**Registrar**: Un registrar è un server che accetta richieste del tipo REGISTER e quindi inserisce le informazioni che riceve nelle richieste del servizio di locazione del dominio che gestisce. Per esempio per il dominio domain.com potrebbe esistere un registrar con opportune regole per la registrazione dei suoi membri, controllo login e password. SIP offre inoltre un servizio di "ricerca utenti" o discovery. Se un utente vuole iniziare una sessione con un altro utente, il registrar deve sapere dove può trovare l'host che gestisce l'utente chiamato.
La discovery è compiuta da elementi di reti SIP come proxy server e server per la redirezione (redirect) che sono responsabili, dopo la ricezione di una richiesta, di determinare dove essa va inviata, basandosi sulla conoscenza che hanno della locazione dell'utente. Per fare ciò i vari dispositivi supportanti SIP consultano un servizio astratto chiamato servizio di locazione il quale fornisce il legame utente indirizzo ip o il riferimento ad altre URI, in modo analogo al funzionamento dei sistemi DNS di Internet.

**Proxy, Proxy Server**: Un proxy è una entità intermedia che agisce sia da server che da client allo scopo di fare richieste per conto di altri client. Un server proxy fondamentalmente ritrasmette in maniera intelligente le richieste, inviando ogni richiesta ad un'entità più vicina all'utente destinatario del messaggio. I proxy possono essere utilizzati anche per fissare determinati criteri di protezione come, ad esempio, l'obbligo per un determinato utente di autenticarsi prima di effettuare una chiamata.

Un proxy interpreta e se necessario riscrive parti specifiche di una richiesta prima di inoltrarla. I proxy SIP sono quindi componenti che indirizzano le richieste SIP verso user agent e le risposte verso i client user agent. Una richiesta può attraversare più proxy prima di raggiungere lo UAS finale. Ognuno di questi prenderà delle decisioni, una volta arrivato

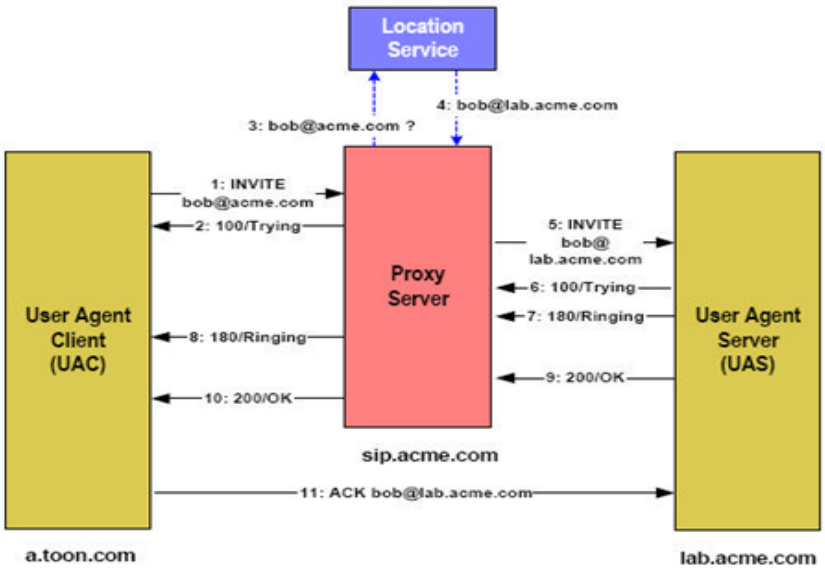


Figura 3.9: Proxy SIP

all'UAS finale la risposta ripercorrerà il percorso inverso della richiesta fino ad arrivare all'UAS originante.
Risulta particolarmente utile legare un proxy ad un particolare dominio e proprio per questo motivo la maggior parte dei dispositivi con supporto SIP è in grado di tradurre i nomi dei domini attraverso l'accesso a server DNS.
**Redirect Server**: Un server per la redirezione è uno user agent server che genera responsi di tipo 3xx. Questi indicano di contattare un differente insieme di URI. Nei casi in cui il carico su un particolare proxy sia troppo alto può essere preferibile ridurre il carico inviando le richieste a differenti proxy. Il processo è simile a quello visibile nel web in cui un server centrale offre semplicemente la redirezione ad altri server aumentando così la qualità del servizio.

La redirezione consente ad un server di dare informazioni utili allo UA (User Agent) ma al contempo di far cadere la nuova connessione aperta con il client. Si potrebbe decidere ad esempio che dopo un certo numero di connessioni un proxy, raggiunto il carico considerato massimo, redirezioni tutte le richieste verso altri proxy. Infatti quando l'UA originante riceve la risposta invierà nuovamente la richiesta ad un altro server.

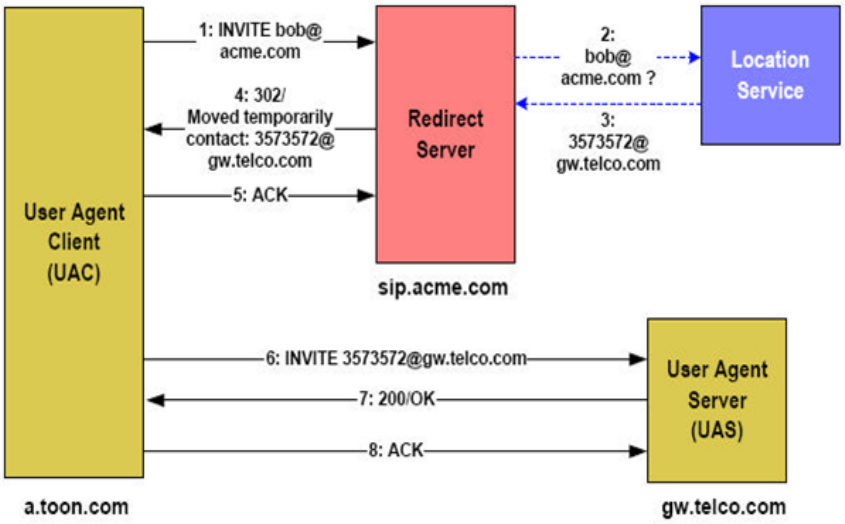


Figura 3.10: Redirect Server

### 3.3.5 Gestione di una chiamata

Al momento della generazione di una chiamata, eseguita ad esempio con un componente push to talk, l'UA residente sul computer genererà un messaggio di tipo Invite verso il proxy a cui è registrato. Se l'indirizzo, nella forma user@domain.ext, non è riconosciuto dal proxy (non è fra gli UA a lui registrati localmente), il messaggio verrà inoltrato ai proxy di livello superiore finché non sarà verificata la presenza del UA user: a questo punto sarà effettuato l'invito per la comunicazione che potrà avvenire nelle seguenti modalità:

1. L'utente ricevente può contattare direttamente l'UA che ha effettuato la chiamata utilizzando il suo indirizzo, grazie ai campi contenuti nel messaggio SIP.

2. L'utente ricevente può accettare la chiamata inviando un messaggio di riscontro al suo proxy server. La conferma seguirà il percorso inverso a quello dell'andata. Una volta che la conferma avrà raggiunto il mittente della richiesta la comunicazione vocale potrà iniziare.

Nel caso che l'UA sia mobile esso deve registrarsi per segnalare la sua posizione presso un registrar che la manterrà presso un Location Server.

Quest'ultimo sarà utilizzato da qualsiasi proxy per ottenere la posizione attuale dell'UA.

## 3.4 SIP vs H.323

E' facile intuire come SIP ed H.323 abbiano sia dei vantaggi che degli svantaggi evidenti. Il principale pregio di H.323 è di essere stabile e collaudato ma la sua realizzazione risulta eccessivamente costosa. E' stato infatti scelto da aziende di telecomunicazioni per sostituire linee telefoniche tradizionali su larga scala. H.323 inoltre offre delle specifiche per la realizzazione dei gateway che sono già presenti in commercio. Sarebbe per lo meno ambizioso gestire lo stesso sistema con SIP fino a quando questo protocollo non sarà completo e supportato interamente da ogni terminale. Infatti, eventuali gateway in grado di interfacciarsi con linee esistenti devono essere ancora realizzati.
D'altronde H.323 ha una struttura più rigida e non si presta a sviluppi futuri o ad altri utilizzi, al contrario di SIP. In generale H.323 essendosi diffuso prima ha trovato mercato nella sostituzione di centralini analogici, nella gestione delle videoconferenze a livello aziendale e ovunque ci fosse bisogno di un servizio affidabile. SIP ha invece offerto una possibilità a molti sviluppatori e produttori di hardware di affacciarsi sul mondo del VoIP con un protocollo semplice da gestire e da controllare.
Dal punto di vista progettuale si può quindi affermare che i due protocolli lentamente stanno raggiungendo una complessità paragonabile e si stanno dimostrando ugualmente validi. SIP infatti nella sua architettura possiede entità simili a quelle presenti in H.323 per gestire più UA in maniera ordinata mentre H.323 ha aggiunto la possibilità di gestire altri contenuti. Esistono quindi notevoli differenze tra i due protocolli che è utile evidenziare.

**Architettura** La struttura di SIP può essere definita modulare. SIP si occupa direttamente solo della segnalazione delle chiamate, della locazione degli utenti e della registrazione degli UA mentre altre funzionalità come QoS, directory access, service discovery risiedono in protocolli ortogonali.
H.323 al contrario, utilizzando protocolli esistenti, ha creato una soluzione monolitica. In un sistema utilizzante H.323 tutto risulta già integrato.

**Interoperabilità** Un'altra differenza tra SIP e H.323 è la interoperabilità tra le diverse versioni. Ogni nuova versione di H.323 mantiene infatti la compatibilità con le precedenti versioni, sono state previste inoltre delle direttive per lo sviluppo di tale modifiche. SIP invece durante il suo sviluppo ha abbandonato alcune convenzioni tipiche delle prime versioni del protocollo.
Questo approccio rappresenta in parte un vantaggio poiché errori progettuali possono essere risolti definitivamente, ma compromette la compatibilità con le precedenti versioni. E' ancora da notare che le relazioni tra le varie componenti per il controllo di un sistema H.323 sono molto strette.
Questo significa che ogni modifica delle specifiche va accuratamente testata e concordata anticipatamente mentre SIP, utilizzando una sola porta per il controllo e le segnalazioni delle trasmissioni, è più facilmente controllabile.

**Codifica dei messaggi** H.323 utilizza una codifica binaria dei messaggi, questa è preferibile perché sicuramente più compressa di un messaggio intelligibile come nel caso di SIP. Questo si traduce in un consumo inferiore di banda che viene però in genere nascosto dal traffico voce il quale occupa la porzione della banda di gran lunga più elevata. Il vantaggio di utilizzare testo ascii per i messaggi sta nella facilità di modifica del pr otocollo e nel suo debugging poiché il contenuto è immediatamente comprensibile. Inoltre possono essere inclusi altri tipi di informazioni come URL in maniera rapida. Chiaramente occorre prevedere una discreta potenza di calcolo per effettuare il parsing del messaggio su ogni dispositivo.

**Protocollo di trasporto** Sia SIP che H.323 possono utilizzare TCP oppure UDP come protocollo di trasporto. Con H.323 viene utilizzato TCP per la segnalazione della chiamata e il controllo della stessa per la criticità dei ruoli che ricoprono. RAS invece, che viene utilizzato per le richieste di un terminale al gatekeeper, utilizza UDP poiché la registrazione ad un servizio può essere in genere ripetuta e perchè l'utilizzo di UDP permette di invii broadcast per l'individuazione della locazione del gatekeeper. SIP richiede una sola porta per il controllo ed in genere viene scelto UDP per il minore overhead.

**Indirizzi** H.323 è stato realizzato per integrarsi nell'architettura telefonica esistente e quindi supporta sia i numeri E.164 (es. +39 011 72727272) con cui vengono gestiti tutti i numeri delle linee urbane in tutto il mondo e sia le URL. Il supporto per i numeri E.164 è necessario per implementare gateway che possano comunicare con linee fisse. H.323 però era stato realizzato con in mente la comunicazione all'interno di una LAN e quindi la gestione delle URL è stata aggiunta in seguito. SIP invece è stato creato specificatamente per inserirsi con efficacia all'interno dell'insieme di applicazioni Internet esistenti e quindi le destinazioni di una chiamata sono solo URL.
Riguardo all'incapacità di SIP di interpretare numeri telefonici internazionali su linee PSTN può tornare in aiuto il fatto che i numeri E.164 possono essere utilizzati per ottenere il nome di un dominio utilizzando Enum.
Enum alloca una specifica zona, e164.arpa per utilizzare numeri E.164. Un qualsiasi numero telefonico, come +39 0461 809912 può essere trasformato nel nome di un host invertendo i numeri, separandoli con punti e aggiungendo il suffisso e164.arpa, in questo modo: 2.1.9.9.0.8.1.6.4.9.3.e164.arpa. L'hostname può quindi essere utilizzato per verificare se esiste un indirizzo associato che potrà essere l'indirizzo di un proxy SIP o in alternativa di un terminale SIP. In questo modo un numero telefonico che rispetta la raccomandazione e164 può essere utilizzato per risalire ad indirizzi IP e quindi a possibili telefoni VoIP.

**Funzionamento con linee PSTN** All'interno delle raccomandazioni H.323 sono presenti specifiche direttive per la comunicazione con linee analogiche. H.323 è stato realizzando prendendo come esempio la tecnologia PSTN e protocolli come Q.931 che gestiscono il controllo della chiamata sulle linee ISDN PRI e BRI. H.323 gestisce attraverso i media gateway la comunicazione con linee tradizionali, nonostante sia completamente basato sullo scambio di pacchetti. Il protocollo SIP invece non prevede nulla del genere per il semplice fatto che non si occupa di specifiche hardware ma di stabilire un protocollo per creare delle sessioni. E' comunque possibile implementare dei gateway ma non sono rilasciate specifiche per la realizzazione di tali apparati.

**Conferenza video e dati** In questo caso H.323 supporta sia la conferenza dati che video. Esistono procedure per fornire controllo sia alla conferenza che la sincronizzazione tra audio e video. SIP al momento

non prevede nulla riguardo la sincronizzazione.

**Utilizzo risorse** Sia SIP che H.323 si basano su protocolli già esistenti per lo scambio dei frame dati veri e propri e cioè RTP e RTCP. In questo modo tre porte UDP sono già utilizzate, nel caso di SIP un'altra porta è necessaria per il controllo e lo stesso si può dire di H.323 ma solo se la comunicazione tra due telefoni è diretta e questa è già stata inizializzata. Sfortunatamente questo caso si presenta solo in configurazioni molto semplici, non appena vengono coinvolti gatekeeper e gateway il numero di porte richieste per gestire il controllo della comunicazione sale a tre.

# Capitolo 4

# Minacce ai Sistemi Voip

## 4.1 Introduzione

Le installazioni VoIP possono essere oggetto di attacchi di diverso tipo. Ciò è dovuto principalmente al fatto che all'utente finale deve essere esposto un gran numero di interfacce e di protocolli, al fatto che la qualità del servizio della rete è un fattore determinante per la qualità del sistema VoIP, ma anche perché, di solito, l'infrastruttura è piuttosto complessa.

In questo capitolo discuteremo di quelli che sono i tipi di attacco più comuni per un sistema VoIP, dando una panoramica generale delle tecniche utilizzate per difendersi da essi.

## 4.2 Tipi di attacco più comuni

La tecnica di attacco più semplice, benché non sempre fruttuosa, è il DoS: semplice da portare a termine, generalmente anonimo e molto efficace. Per esempio, si potrebbe sottoporre a un attacco DoS una certa infrastruttura mediante l'inoltro di un traffico elevato di false impostazioni di chiamata (SIP INVITE), oppure una singola apparecchiatura telefonica "sommergendola" di traffico unicast o multicast.
Qualsiasi attacco di rete di tipo DoS, intenzionale o causato da un worm, avrà comunque un effetto negativo sulla qualità del sistema VoIP, anche se la rete non è assoggettabile a QoS. Il cosiddetto call spoofing, ossia l'impersonificazione dell'identità di un chiamante, è un altro meccanismo di attacco assai diffuso: consiste nel rubare l'identificativo CLID (Caller ID) in fase di chiamata. Questa tecnica potrebbe permettere anche l'accesso alla casella vocale dell'utente legittimo, se il sistema si basa soltanto

sul CLID e non richiede un PIN di identificazione.
L'introduzione di dati (injection) in una comunicazione in corso è un'altra strada percorribile, ma piuttosto complessa, e i risultati potrebbero non essere perfetti, in quanto gli interlocutori potrebbero rilevarla. Questa tecnica richiede l´iniezione di pacchetti RTP, ma alcuni stack TCP/IP sui sistemi intermedi (gateway) o estremi (apparecchiature telefoniche o software) della comunicazione, potrebbero essere soggetti a strani comportamenti al ricevimento di dati RTP fuori sequenza o quasi duplicati, provocando l'interruzione della comunicazione.
Alterare la configurazione del dispositivo telefonico è un altro metodo semplice. Se il telefono è connesso in rete, si potrebbe cercare di accedervi tramite le comuni interfacce di gestione che potrebbero essere esposte, per esempio un'interfaccia telnet CLI o HTTP non protetta, con una password facile o addirittura assente.
Se questo accesso non è possibile, si potrebbe cercare di prendere possesso dell'apparecchiatura o del software ricorrendo ai server DHCP e TFTP: all'avvio, l'apparecchiatura ottiene le informazioni sull'indirizzo IP e sulla rete tramite DHCP, quindi scarica la propria configurazione e a volte il firmware aggiornato via TFTP. A seconda che il sistema si basi soltanto sull'indirizzo IP o meno, il protocollo DNS potrebbe essere associato al processo e renderne possibile lo spoofing.
La maggior parte delle nuove applicazioni legate a VoIP, come i sistemi di posta vocale avanzati, la messaggistica istantanea, i servizi di calendario o la gestione degli utenti, si basa sui servizi web. Queste applicazioni sono spesso ricche di falle, quali il cross-site scripting, alcuni JavaScript usati per la verifica dei form sul lato client, l'iniezione SQL, e così via: pertanto non è escluso che i normali metodi per violare le applicazioni web possano rivelarsi utili per l'accesso al sistema e l'enumerazione di numeri importanti, caselle vocali utente, CRD (Call Detail Records) e via dicendo. In alcune situazioni, per esempio quando occorre registrare le chiamate per motivi legali, l'ottenimento dell'accesso al sistema di registrazione delle chiamate potrebbe consentire di impadronirsi di informazioni confidenziali. Naturalmente, per molti hacker si tratta di un'occasione preziosa, per cui è essenziale implementare soluzioni di sicurezza basate sugli host.
Il rilevamento delle frodi è un altro problema, benché interessi soprattutto i fornitori di telefonia: agli utenti non dovrebbe essere permesso accedere a numeri riservati. Ciò è particolarmente importante per le aziende che mettono a disposizione i gateway VoIP-to-PSTN.
Inoltre bisogna tenere presente che, come nel caso della posta elettronica, la mancata eliminazione dell'header SIP potrebbe fornire all'hacker informazioni sulla struttura della rete e di altro tipo.

## 4.3 Un'approccio integrato per la sicurezza

A fronte di vantaggi riconosciuti, evidenti ed innegabili, l'adozione di soluzioni di Telefonia IP non può prescindere da considerazioni di buonsenso mirate a proteggere funzionalità e riservatezza delle comunicazioni: ecco perchè nella scelta di sistemi ed architetture è opportuno optare per quelle soluzioni in grado di integrare nativamente funzionalità di sicurezza non solo all'interno dei singoli apparati, ma anche nell'interoperabilità con il "Sistema IP" nel suo complesso.

Uno dei principali fattori di spinta nell'evoluzione del Networking di questi anni è la cosiddetta "convergenza", ovvero la visione di una singola piattaforma di comunicazione aperta per il trasporto di voce, video e dati, che abiliti e semplifichi le interazioni tra singoli utenti ed organizzazioni indipendentemente dagli strumenti utilizzati o dal luogo in cui ci si trovi. In effetti si tratta di uno dei più importanti trend evolutivi nel settore delle telecomunicazioni che promette una serie di implicazioni di notevole portata nel prossimo futuro, ma da cui già oggi è possibile trarre tangibile vantaggio.

Ad esempio:

- abilitazione di servizi innovativi, a beneficio di nuovi livelli di produttività e flessibilità delle organizzazioni

- consolidamento delle infrastrutture di comunicazione

- riduzione dei costi operativi di gestione e manutenzione

Un fondamentale elemento costitutivo di questo scenario è certamente rappresentato dalla Telefonia su IP, o IP Telephony. Il principio di funzionamento dell'IP Telephony consiste nel digitalizzare in tempo reale il contenuto delle conversazioni telefoniche trasformandolo in un flusso di pacchetti dati che insieme ai relativi messaggi di segnalazione viene appunto instradato e gestito sulla stessa infrastruttura di rete IP utilizzata per il trasporto dei dati. In effetti qualunque organizzazione si trovi oggi a sostituire il proprio centralino tradizionale non può prescindere dal prendere quantomeno in considerazione questa tecnologia. Ma è altrettando interessante il fatto che per quest'area del comparto ICT è prevista una

consistente e continua crescita vicina al 50% anche per i prossimi anni, allorchè un numero sempre maggiore di aziende ed organizzazioni coglierà i vantaggi della riduzione dei costi operativi associati a questa tecnologia.

Eccone alcuni esempi:

- riduzione dell'investimento altrimenti necessario per realizzare due reti fisicamente e tecnologicamente distinte, una per la voce e una per i dati;

- minori costi associati alla maggior semplicità di gestione per le modifiche delle utenze e l'espansione di nuove sedi o nuovi utenti;

- economie di scala nella gestione centralizzata delle chiamate, anche di sedi remote

- riduzione dei costi e dei tempi legati all'introduzione di nuovi servizi e applicazioni, su una piattaforma IP nativamente predisposta, invece che su sistemi PBX proprietari.

Se dunque i benefici applicativi di questa nuova tecnologia di comunicazione sono più che evidenti, è importante acquisire altrettanta consapevolezza delle implicazioni di sicurezza legate all'introduzione della Telefonia IP nelle reti aziendali. Se da un lato l'introduzione di queste nuove funzionalità e tecnologie rappresenta una notevole opportunità, dall'altro introduce nuove complessità e rischi di sicurezza che è naturalmente opportuno comprendere per indirizzare proattivamente. La disponibilità è la prima considerazione da fare, direttamente legata al fatto che la convergenza di voce e dati comporta che sulla stessa infrastruttura di telecomunicazioni siano condensate due reti storicamente disgiunte. Ad oggi le reti dati possono avere eccellenti livelli di ridondanza e disponibilità, anche a beneficio delle applicazioni di Telefonia IP. Seppure le due reti convergano fisicamente sulla stessa infrastruttura, tuttavia queste restano logicamente separate e proprio la separazione virtuale fra le due realtà è una considerazione estremamente importante.

### 4.3.1 VLAN

Grazie all'utilizzo di meccanismi tipici delle reti IP come le Virtual LAN (VLAN) e la separazione di Livello 2 si ha la possibilità di attestare sulla stessa infrastruttura fisica le due applicazioni attribuendo politiche di traffico e di sicurezza differenziate, come è appunto opportuno e necessario date proprio le diverse peculiarità dei due tipi di traffico in questione: da una parte il traffico dati, a cui va dato un certo livello di priorità anche in dipendenza delle specifiche applicazioni, e dall'altra il traffico voce, che invece deve viaggiare a massima priorità trattandosi di un'applicazione real-time, in cui parametri trasmissivi di latenza e jitter sono assolutamente critici. Un ulteriore requisito è che il traffico voce, che può viaggiare in chiaro, resti virtualmente separato dal segmento dove viaggiano normali dati e dove l'utenza è attestata, minimizzando così la possibilità di attacchi di Denial of Service scatenati utilizzando la potenza elaborativa dei PC. Ad oggi, l'utilizzo di VLAN offre un ottimo livello di sicurezza per la separazione dei dati, grazie alle ormai mature tecnologie di livello 2 per la prevenzione di attacchi. Il che offre quindi tutte le possibilità per prevenire efficacemente gli attacchi mirati a violare la separazione virtuale fra i due mondi, che potrebbero concretizzarsi sotto forma di furto di QoS, eavesdropping (intercettazione di chiamate), sfruttamento di Covert Channels e disservizi nella segnalazione, che possono portare dalla indebita redirezione delle chiamate (call hijacking) fino a parziale e totale disservizio.

### 4.3.2 Cifratura

Un'altra importante considerazione riguarda la confidenzialità del traffico voce. Infatti grazie all'ampia disponibilità di sniffer del traffico, la trasmissione di pacchetti voce in chiaro su una rete IP comporta inevitabilmente possibili implicazioni di intercettazione telefonica a chiunque abbia accesso alla rete. Oltre alla separazione virtuale fra segmento voce e dati, che per le ragioni già illustrate contribuisce a minimizzare questo rischio, un'altra opzione disponibile è la cifratura del traffico voce. Ciò ovviamente introduce tempi di latenza che vanno opportunamente considerati in fase di progetto, insieme al relativo impatto sulla garanzia della Qualità del Servizio (QoS), che va conseguentemente riconsiderata. La QoS non è naturalmente un requisito a cui è sufficiente trovare corrispondenza a livello di singolo dispositivo, quanto piuttosto lungo l'intero percorso che un flusso di traffico voce su IP si trova a seguire, affinchè alle due estremità di una sessione possa essere effettivamente ricostruita la

necessaria qualità. È proprio questo requisito di supporto "End-to-End" di tutte le funzionalità e meccanismi di gestione del traffico voce su IP che entra pesantemente in gioco anche quando vengono presi in considerazione gli aspetti di Sicurezza di una realizzazione IP Telephony. Tutto questo va infatti garantito anche laddove il traffico voce venga cifrato, ed anche in questo caso è possibile farlo attraverso opportuni meccanismi grazie ai quali sebbene il flusso di pacchetti voce si trovi a viaggiare cifrato all'interno di un tunnel IPSec, continui ad essere garantita la priorità di instradamento rispetto agli altri flussi.

### 4.3.3 Monitoraggio

Resta ancora da considerare la vulnerabilità applicativa dei server su cui si basa l'intera gestione dell'applicazione di telefonia IP. Su questo tipo di sistemi il controllo delle chiamate e la segnalazione vengono infatti effettuate da un'applicazione che in molte implementazioni risiede su un sistema operativo general purpose. Di quest'ultimo, ovviamente, l'applicazione di call control rischia di ereditare tutte le eventuali vulnerabilità. Si è affermata quindi la pratica di dotare questo genere di server di sistemi antivirus aggiornati per intercettare eventuali virus o worms già conosciuti eventualmente in circolazione, ma solo le soluzioni più evolute prevedono uno strato software con funzionalità tipiche di host intrusion protection (H-IPS) e personal firewall, in modo da costruire uno scudo virtuale a garanzia della stabilità, integrità e disponibilità di server così critici. Il server di gestione delle chiamate viene così protetto e configurato per eseguire solo le operazioni strettamente legate all'applicazione specifica di gestione del sistema di telefonia IP, disattivando così tutti gli ulteriori servizi non necessari. Con tale tecnica integrata in un sistema di telefonia IP aziendale è possibile minimizzare i pericoli potenzialmente derivanti dall'utilizzo di sistemi operativi generici come piattaforma per la gestione delle chiamate, che potrebbero altrimenti risultare vulnerabili a worm, virus e attacchi di tipo Denial-of-Service (DoS).

### 4.3.4 Filtraggio

Anche il filtraggio del traffico voce nel momento in cui questo attraversa un firewall è un aspetto molto importante, e considerato critico e difficoltoso fino a poco tempo fa proprio per l'utilizzo dei nuovi protocolli multimediali di segnalazione e trasporto (SIP, H.323, SCCP, MGCP, Megaco/H.248) introdotti sulle reti IP dalle applicazioni voce. In realtà

con l'ultima generazione di firewall grazie alla loro potenza ed elevata intelligenza di sistema, l'accoppiata VoIP/Firewall oggi non rappresenta più un elemento di criticità. Per le ragioni su esposte questo aspetto risulta di particolare importanza nel qualificare un firewall rispetto ai suoi livelli di integrazione nativa con una soluzione di telefonia IP. Altre considerazioni riguardanti i sistemi di telefonia IP e le loro implicazioni di sicurezza possono poi essere ricondotte a più generali best practice per il miglior contenimento di rischi su di una rete IP, trattandosi ovviamente di indicazioni e pratiche di utilizzo che finiscono per andare a beneficio della soluzione IP Telephony, oltre che della rete su cui questa si basa.

## 4.4 Conclusioni

Tutta la discussione precedente concorre a definire un quadro complessivo in cui è quantomai indispensabile che l'approccio alla sicurezza delle implementazioni IP Telephony faccia riferimento alle più aggiornate ed autorevoli best practice relative al progetto di reti IP convergenti intrinsecamente sicure. Da questo punto di vista è interessante notare che parallelamente alla diffusione dell'IP Telephony si stanno rendendo disponibili anche raccomandazioni interessanti ed autorevoli e metodologie relative alle più opportune considerazioni di sicurezza da effettuare in fase di progetto, implementazione e utilizzo di queste nuove tecnologie. Ad esempio già da diverso tempo il NIST (National Institute of Standards and Technology: http://www.nist.gov) ha pubblicato un documento di riferimento intitolato "Security Considerations for Voice Over IP Systems", e sono stati pubblicati i risultati di test importanti ed autorevoli e di prove comparative sulla sicurezza di soluzioni IP Telephony.

In sintesi, questi i criteri a cui in generale si riferiscono queste linee guida:

- protezione globale non solo dei singoli dispositivi o sistemi, combinando diverse funzionalità e componenti di sicurezza, e agendo su fronti diversi e complementari garantendo quindi i desiderati livelli di protezione, controllo e gestione del rischio;

- stretta integrazione, supporto ed interoperabilità end-to-end delle funzionalità IP Telephony tra le componenti di sicurezza e la stessa infrastruttura di rete IP;

- architetture di sicurezza multilivello, in modo da garantire il contenimento delle minacce a garanzia dell'intera infrastruttura nell'evenienza in cui un singolo sistema venga compromesso.

L'attuazione di queste linee guida per la messa in sicurezza di sistemi di telefonia IP porta quindi alla definizione delle tre componenti chiave di un progetto di IP Telephony sicuro:

1. Elementi di sicurezza perimetrale, per la segregazione di opportune sezioni di rete tramite la quale sia possibile controllare che solo i legittimi dispositivi e applicazioni abbiano accesso a risorse particolarmente critiche, come un CallManager;

2. Elementi per la risevatezza delle comunicazioni e la sicurezza delle connessioni, tramite meccanismi di autenticazione e opportuna segmentazione del traffico dati rispetto al traffico voce in ambito locale (LAN), e l'impiego di VPN Virtual Private Network con supporto della Quality of Service per il traffico voce in ambito geografico (WAN). A questo proposito l'utilizzo di certificati digitali su telefoni IP e piattaforme di call processing, unito alla cifratura di segnalazione e traffico voce, è strumento qualificante a garanzia della protezione del sistema di telefonia IP nel suo complesso.

3. Elementi di Intrusion Protection sia sugli specifici segmenti di rete interessati, sia sui server critici su cui risiedono le applicazioni Voce , ai quali è affidata l'analisi in tempo reale del contenuto e del contesto dei singoli pacchetti per controllare tentativi di intrusione o attività sospette.

Le considerazioni sin qui esposte portano dunque alla logica conclusione che l'integrazione sicura di applicazioni voce e dati su un'unica infrastruttura convergente di trasporto è una questione abbastanza complessa, che richiede un impegno maggiore e ancora più competente rispetto alla sola sicurezza delle reti dati. Qualunque strategia premiante per la sicurezza dei sistemi IP Telephony non può infatti prescindere da un più ampio approccio al disegno delle moderne reti convergenti su IP: un approccio in cui Sicurezza, funzionalità di rete e nuove tecnologie sono nativamente integrati.

# Capitolo 5

# Denial of Service

## 5.1 Introduzione

Gli attacchi Denial-of-Service possono influenzare qualsiasi servizio di rete basato su IP. L'impatto di un attacco DoS può variare da una semplice degradazione del servizio offerto ad una completa distruzione dello stesso. Ci sono diverse classi di attacchi DoS. Un tipo di attacco in cui i pacchetti sono semplicemente inondati nella rete target, da parte di sorgenti esterne multiple, è chiamato attacco distributed denial-of-service (DDoS). Le condizioni della seconda grande classe di attacchi DoS si verificano quanto i dispositivi della rete interna sono gli obiettivi di un grande flusso di pacchetti proveniente dall'interno della rete stessa. Tale flusso non è gestibile e causa il fallimento del dispositivo stesso.

Nè le verifiche di integrità, nè la crittografia possono prevenire questo tipo di attacco. Gli attacchi DoS o DDoS sono caratterizzati semplicemente dal volume dei pacchetti inviati verso il computer della vittima; se questi pacchetti sono firmati da un server, se contengono indirizzi IP reali o falsi oppure se sono crittografati con una chiave fittizia, nessuna di queste caratteristiche è rilevanti ai fini dell'attacco. É difficile difendersi dagli attacchi DoS, e poichè VoIP è solamente un altro servizio di rete basato su IP, anch'esso è suscettibile agli attacchi DoS così come qualsiasi altro servizio basato su IP.

In questo capitolo parleremo degli attacchi DoS, in quanto sono particolarmente efficienti sui servizi VoIP ed altri servizi Real-Time, dal momento che tali servizi sono molto sensibili alle condizioni della rete.

## 5.2 Attacchi portati da un singolo host

Questi tipi di attacco, provenendo da un'unica fonte, sono potenzialmente rintracciabili.

### 5.2.1 Syn-Flood

Storicamente il Syn-Flooding rappresenta il capostipite degli attacchi DoS, che trova le sue dirette radici nel Ping of Death. Col termine Syn Flooding, letteralmente tradotto con inondazione di pacchetti di tipo Syn, nasce dal fatto che tutte le volte che un utente fa click su di un link di una pagina web richiede l'apertura di una connessione (di tipo TCP) verso quel sito; questo avviene seguendo una serie di passi, il primo dei quali consiste nell'invio di un pacchetto TCP che richiede l'apertura di una connessione.

Tutte le regole di funzionamento del protocollo TCP esigono che il sistema risponda allocando alcune risorse (in pratica memoria) per la connessione. Se si programma opportunamente un semplice PC, è possibile richiedere l'apertura di diverse migliaia di connessioni al secondo, che inondando il server, ne consumano rapidamente tutta la memoria, bloccandolo o mandandolo in crash.

Un esempio potrebbe essere il seguente: l'attaccante, identificato dal nome STE, invia una serie di richieste alla sua vittima, identificata col nome CRI: la macchina server, sulla quale vengono eseguiti dei servizi, non sarà in grado di gestire tutte le richieste e i servizi stessi andranno in crash, risultando prima molto rallentati e poi, successivamente, inaccessibili. In questa maniera, un utente qualunque (identificato dal nome UTENTE) non sarà in grado di accedere ai servizi, ricevendo un errore di richiesta scaduta o timeout.

L'attacco Syn-Flood usa strumenti che rientrano nella categoria Tribe Flood Network (TFN) ed agisce creando delle connessioni che si rivelano aperte a metà.

Il protocollo classico usato nei DoS è il ping, inviandone a milioni si riuscirà a bloccare l'operatività di qualunque sito Internet, ma trattandosi di un modello di attacco uno a uno, ad un pacchetto in uscita corrisponderà la ricezione di un solo pacchetto al sistema attaccato.

Occorrerà quindi che i cracker possano disporre di un gran numero di PC client, controllati, ma non è così facile inoculare il codice maligno in un numero tanto elevato di macchine grazie all'azione specifica di antivirus, patch di sicurezza e tecnici informatici.

### 5.2.2 Smurf

Una modalità di attacco più sofisticata, detta Smurf attack, utilizza un flusso di pacchetti modesto, in grado di passare attraverso una normale connessione via modem, ed una rete esterna, che sia stata mal configurata, che agisce da moltiplicatore di pacchetti, i quali si dirigono infine verso il bersaglio finale lungo linee di comunicazione ad alta velocità.

Tecnicamente, viene mandato uno o più pacchetti di broadcast verso una rete esterna composta da un numero maggiore possibile di host e con l'indirizzo mittente che punta al bersaglio (broadcast storm).

Ad esempio può venir usata una richiesta echo ICMP (Internet Control Message Protocol) precedentemente falsificata da chi attua materialmente l'attacco informatico.

Tuttavia questo tipo di attacco è possibile solo in presenza di reti che abbiano grossolani errori di configurazione dei sistemi (detti router) che le collegano tra loro e con Internet.

## 5.3 Attacchi provenienti da più host

In questi attacchi il bersaglio viene attaccato contemporaneamente da più fonti, rendendo difficile rintracciare l'attaccante originario.

### 5.3.1 DDoS

Una variante dell'attacco proveniente da un singolo host è il DDoS (Distributed Denial of Service), che ha un funzionamento identico ma realizzato utilizzando numerose macchine attaccanti che insieme costituiscono una botnet.

Gli attaccanti tendono a non esporsi direttamente, dato che per le forze dell'ordine sarebbe relativamente semplice risalire ai computer utilizzati per l'attacco. Gli attaccanti, per evitare di essere individuati e per avere a disposizione un numero sufficiente di computer per l'attacco, inizialmente infettano un numero elevato di computer con dei virus o worm che lasciano aperte delle backdoor a loro riservate. I computer che sono controllati dall'attaccante vengono chiamati zombie.

Tutti i computer infettati entrano a far parte di una botnet, a libera disposizione dell'attaccante: una nota interessante è data dalla distinzione tra quelle che sono le macchine che eseguono un Sistema Operativo Windows (definiti, in gergo, rxbot) e quelle che invece eseguono un

sistema Unix, particolarmente adatte all'UDP Flooding (Flooding sul protocollo UDP).

Una particolarità degli zombies Windows è data dalla possibilità, per l'attaccante, di programmare un trojan in grado di diffondersi automaticamente a tutta una serie di contatti presenti sul computer infettato (definita, in gergo, funzione di auto-spreading): contatti contenuti nella rubrica degli indirizzi e nei contatti di programmi di Instant Messaging, come Microsoft Messenger, permettendo così al computer zombie di infettare, in maniera completamente autonoma, altre macchine che, a loro volta, diverranno parte della botnet dell'attaccante.

Quando il numero di zombies è ritenuto adeguato, o quando viene a verificarsi una data condizione, i computer infetti si attivano e sommergono il server bersaglio di richieste di connessione. Con l'avvento della banda larga il fenomeno dei DDOS sta assumendo proporzioni preoccupanti, dato che attualmente esistono milioni di persone dotate di una connessione ad Internet molto veloce e permanente ma con scarse o nulle conoscenze e contromisure riguardanti la sicurezza informatica.

Il danno maggiore dell'attacco di tipo DDoS è dovuto principalmente alla asimmetria che si viene a creare tra la richiesta e le risposte correlate in una sessione DNS (Domain Name System). Il flusso enorme di risposte generato provocheranno nel sistema una tale inondazione di traffico rendendo il server inadeguato alla gestione delle abituali funzioni on-line.

Inoltrando, al Sito preso di mira, una risposta di alcuni Kilobyte, per ogni richiesta contenente solo pochi bytes, si ottiene un'amplificazione esponenziale tale da saturare i canali dati più capienti, raggiungendo con il DDoS livelli finora inattuabili con gli altri tipi di attacco DoS.

Le configurazioni predefinite, standard e quelle consigliate di Firewall si rivelano utili a contrastare solo gli attacchi sferrati dall'esterno, ad esempio di un'azienda, ma poiché il traffico in Rete gestito tramite sistema DNS è vitale, per fronteggiare questo tipo di attacco non si potranno attuare le stesse strategie impiegate nei confronti degli attacchi ai Ping.

Quindi il Network manager dovrà tenere scrupolosamente sotto controllo e monitoraggio i canali di flusso dati e, per escludere l'intervento o contrastare l'azione di un cracker, riconfigurerà il DNS responsabile del sito.

### 5.3.2 DRDoS

Una particolare categoria di DDoS è il cosiddetto Distributed Reflection Denial of Service (DRDoS). In questa particolare tipologia di attacco, il computer attaccante produce delle richieste di connessione

verso server con connessioni di rete molto veloci utilizzando come indirizzo di provenienza non il proprio bensí quello del bersaglio dell'attacco. In questo modo i server risponderanno affermativamente alla richiesta di connessione non all'attaccante ma al bersaglio dell'attacco. Grazie all'effetto moltiplicatore dato dalle ritrasmissioni dei server contattati, che a fronte della mancanza di risposta da parte del bersaglio dell'attacco (apparentemente l'iniziatore della connessione) provvederanno a ritrasmettere (fino a 3 volte solitamente) il pacchetto immaginandolo disperso, entrando cosi in un circolo vizioso che vede rapidamente esaurirsi le risorse del bersaglio.

Quest'ultimo tipo di attacco è particolarmente subdolo perché, a causa della natura delle risposte, è difficilmente schermabile dall'utente comune: infatti se si filtrassero le risposte dei server verrebbe compromessa la funzionalità stessa della connessione di rete impedendo, di fatto, la ricezione anche delle informazioni desiderate. Le risposte dei server, sollecitate dall'attaccante, sono infatti indistinguibili da quelle generate da una richiesta legittima della vittima. Il problema si sta presentando con maggiore incidenza da quando Microsoft ha deciso di rendere le Raw Sockets, interfaccia di accesso al TCP/IP, facilmente disponibili. Le RAW sockets permettono appunto di cambiare l'indirizzo di provenienza del pacchetto per sostituirlo con quello della vittima, fatto che è strumentale per questo tipo di attacco.

## 5.4 Dispositivi di Protezione da Attacchi DoS e Ddos

**Filtraggio dei Dati in arrivo** Implementando i filtri che presiedono all'ingresso, nei propri router e firewall, dei pacchetti contenenti informazioni sulla provenienza dei dati alterate (cioè spoofed), non si otterrà un arresto dell'attacco DoS ma si potrà ricostruire il flusso di traffico qualificato come malefico in tempi relativamente brevi, per consentire la reazione difensiva degli Internet Service Provider (anti spoofing).

**Limitazione del Traffico** Molti router consentono, attualmente, di limitare la quantità di banda utilizzata per la fornitura di un servizio attraverso il campionamento ed analisi dei pacchetti che vi transitano. In caso di attacco non resterà attiva una quantità di banda sufficiente a provocare un danno cospicuo o a bloccare il flusso legittimo dei dati. Questa limitazione si otterrà ad esempio con l'utilizzazione di una macchina Li-

nux che funga da gateway attraverso un'azione CAR (Commited Access Rate), così si bloccherà un attacco DDoS che usi pacchetti ICMP o TCP SYN, poiché viene considerevolmente limitata la banda utilizzabile da questi.

**Sistemi di riconoscimento delle intrusioni** Attraverso tali sistemi di verifica (Intrusion Detection System) vengono individuati i malintenzionati che comunicano con alcune delle macchine nella propria rete, e scoprire se vengono usate malignamente, come pedine per sferrare l'attacco. In particolare i Network Auditing Tools sono programmi che consentono la verifica e l'analisi della rete aziendale alla ricerca di eventuali agenti in grado di provocare un attacco di tipo DDoS.

# Capitolo 6

# Spiare una rete VoIP

## 6.1 Introduzione

Nel corso del tempo le persone hanno cercato di salvaguardare sempre più la privary delle loro comunicazioni. Uno dei migliori esempi ci fu dato da Giulio Cesare, che inventò un crifario a shift per codificare le comunicazioni militari che inviava al proprio esercito con un messaggero. Dai tempi di Giulio Cesare, il campo della crittografia si è sviluppata notevolmente dal supportare qualsiasi forma di comunicazione, incluso VoIP.

Dal momento che VoIP è soltanto un'altra applicazione dati, ci sono diversi modi per salvaguardare la privacy attraverso i vari strati del modello OSI. Sfortunatamente, ci sono anche un gran numero di modi in cui un attaccante può spiare[1] una conversazione VoIP attaccando ognuno di questi livelli. Accedendo ad un punto appropriato della rete, un attaccante può eseguire un gran numero di attacchi oltre quello di spiare semplicemente le conversazioni.

In questo capitolo discuteremo innanzitutto dei modi in cui un attaccante potrebbe ottenere un accesso non autorizzato alla rete, e successivamente ottenuto tale accesso discuteremo di cosa potrebbe fare per spiare una conversazione.

[1]Spiare una rete VoIP si riferisce al termine inglese *"eavesdropping"*. Con *"eavesdropping"* ci si riferisce all'ascolto di nascosto di una conversazione. Per tale motivo abbiamo preferito utilizzare il termine italiano *"spiare"* invece del termine inglese originale.

## 6.2 Ottenere accesso alla Rete

Prima di poter eseguire un qualsiasi tipo di attacco, l'attaccante deve innanzitutto ottenere il livello di accesso appropriato alla rete, altrimenti non potrebbe ascoltarne il traffico. Quelle seguenti sono alcune delle tecniche più popolari che un attaccante ha a disposizione per ottenere tale accesso.

**Compromettere un nodo della rete** L'accesso ad un elemento di una rete VoIP è abbastanza per spiare le conversazioni che passano attraverso esso. Ad esempio, se un hacker riesce a compromettere un endpoint VoIP (telefono IP, PC con softphone, ecc.), allora sarà in grado di spiare soltanto le conversazioni che passano attraverso esso. Se invece riesce a compromettere uno switch, sarà capace di spiare tutte le conversazioni che passano attraverso esso.

**Compromettere un Telefono** Molti telefoni IP hanno diverse caratteristiche che possono facilitare diversi degli attacchi che abbiamo presentato in precedenza. Un buon esempio è il telefono Snom 320. Esso ha una funzionalità, *PCAP Trace*, che permette, a chiunque abbia accesso all'interfaccia web di amministrazione del telefono, di catturarne tutto il traffico.

**Compromettere uno Switch** Un hacker potrebbe essere in grado di ottenere un accesso amministrativo ad uno switch attraverso la sua interfaccia web o attraverso una console telnet. Alcuni switch hanno l'abilità di supportare la modalità Remote Switched Port Analyzer (RSPAN). La modalià RSPAN permette di copiare tutto il traffico su diverse porte di una particolare VLAN per poterlo monitorare. Questo significa che un hacker potrebbe remotamente riconfigurare uno switch al fine di monitorare il traffico su tutte le sue porte.

**Compromettere un Proxy, un Gateway, un PC/Softphone** La sicurezza di una rete VoIP dipende molto da quanto sono sicuri i livelli si supporto su cui si poggia. Costruire un'applicazione VoIP sicura potrebbe non essere sufficiente se il sistema operativo o il firmware su cui viene eseguita può essere compromesso. Molti gateway, proxy e softphone VoIP, vengono eseguiti su sistemi operativi Windows o Linux. Questi sistemi sono soggetti a numerose vulnerabilità che richiedono un continuo lavoro di patching e aggiornamento. Ci sono una varietà di tool per

sfruttare tali vulnerabilità. Ad esempio uno di tali tool, già fornito con una numerosa lista di exploit pronti all'uso è il Framework Metasploit.

Una volta che un host è stato compromesso, c'è una varietà di backdoor e programmi rootkit che l'hacker può copiare sulla macchine al fine di mantenere un accesso remoto ad essa. Nel caso del VoIP, potrebbe copiare dei tool in grado di registrare il traffico VoIP che fluisce attraverso l'host compromesso.

## 6.3 I Rischi per la Privacy VoIP

I maggiori quattro attacchi per spiare una conversazione VoIP sono: *TFTP configuration file sniffing, number harvesting, call pattern tracking, conversation eavesdropping.* Ognuno di questi attacchi richiede che un attaccante abbia accesso a qualche parte della rete in cui del traffico VoIP attivo (bootup,signaling,media,ecc.) sia in transito. Tale accesso può ossere ottenuto da qualsiasi parte, da un softphone a un'accesso a uno switch.

### 6.3.1 TFTP Configuration File Sniffing

Una volta accesi, molti telefoni IP utilizzano TFTP per effettuare il download del proprio file di configurazione. Spesso tale file di configurazione contiene password che possono essere usate per effettuare una connessione diretta con il telefono (in altre parole, telnet, un'interfaccia web, ecc.) ed amministrarlo. Un attaccante che ascolta il traffico quando il telefono effettua il download di tale file, potrebbe ottenere le password e potenzialmente riconfigurare e controllare il telefono IP.

**Attacco** Riuscire ad ottenere tale file potrebbe essere una cosa abbastanza semplice quanto visionare tutto il traffico sulla porta UDP 69 (la porta di default del servizio TFTP). Esistono una vasta varietà di tool per catturare informazioni di tale tipo. Una volta individuato il file di configurazione, è possibile scaricarlo direttamente dal server TFTP e quindi recuperare da esso informazioni come username e passwords.

**Difesa** Proprio per la natura insicura di TFTP, non ci sono molte opzioni per rendere sicuro il canale di comunicazione. Un opzione sarebbe quella di creare una VLAN separata per la comunicazione tra i telefoni e il server TFTP. In questo modo ci assicuriamo che il server TFTP serve

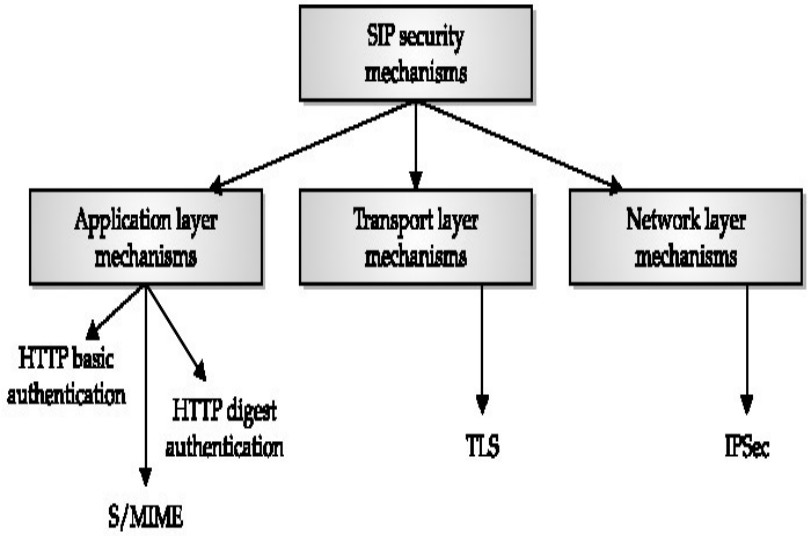


Figura 6.1: Livelli di Sicurezza

soltanto quei telefoni appartenenti alla VLAN. Un altra soluzione sarebbe quella di creare un ACL (Access Control List) sul firewall, in modo da assicurare che solamente indirizzi IP validi siano in grado di accedere al server TFTP.

### 6.3.2 Number Harvesting

Questo attacco descrive il modo in cui un attaccante, monitorando passivamente tutte le chiamate in entrata e in uscita, può costruire un database di numeri di telefono validi. Tale database può essere usato per effettuare attacchi VoIP più avanzati come *signaling manipulation* o *SPIT*.

**Attacco** In ambiente SIP, per eseguire questo tipo di attacco, il modo più semplice è quello di spiare tutto il traffico sulla porta TCP e UDP 5060 ed analizzare i campi dell'header *From:* e *To:*.

**Difesa** Una contromisura è quella di criptare il traffico del protocollo di signaling o al livello di rete (IPSec) o al livello del trasporto (TLS). Inoltre utilizzare VLAN separate potrebbe aiutare ad attenuare il rischio di sniffing sulla rete. La figura 6.1 mostra i vari livelli di sicurezza che possono essere applicati allo stream di signaling tra i vari livelli.

### 6.3.3 Call Pattern Tracking

L'attacco *Call Pattern Tracking* fa un passo avanti rispetto all'attacco *Number Harvesting* in quanto riesce a determinare coloro che stanno comunicando, anche quando la loro conversazione è criptata. Dal punto di vista della legge, questo è un ovvio beneficio in quanto è possibile rilevare eventuali cospirazioni criminali. Tuttavia ci potrebbero essere anche implicazioni nello spionaggio tra aziende, ad esempio qualora un'azienda sia capace di vedere i clienti che vengono contattati dall'aziende concorrenti. Informalmente, tale attacco è simile a rubare il conto mensile del telefono cellulare di qualcuno, allo scopo di vedere tutte le chiamate in entrata e in uscita.

**Attacco** Essenzialmente tale tipo di attacco è simile all'attacco *Numeber Harvesting*, quindi le tecniche di attacco sono le stesse.

**Difesa** Anche le tecniche di difesa sono le stesse citate per l'attacco *Number Harvesting*.

### 6.3.4 Conversation Eavesdropping and Analysis

La paura maggiore di molti utenti VoIP è quella che la propria conversazione possa essere ascoltata da qualche maleintenzionato. Con questo attacco, un maleintenzionato cerca di registrare una o entrambi gli estremi di una comunicazione telefonica. Oltre che ad spiare il contenuto della conversazione, il maleintenzionato può anche utilizzare degli strumenti che gli permettono di decodificare eventuali tasti premuti durante la chiamata. I tasti, conosciuti come *dual-tone multifrequency (DTMF) tones*, sono spesso usati, quando il chiamante fornisce numeri di pin, ad esempio al servizio telefonico del proprio istituito bancario. Essere capaci di catturare tali informazioni potrebbe significare essere capaci di riutilizzarli per avere accesso allo stesso servizio telefonico utilizzato dalla vittima.

**Attacco** Una volta ottenuto l'accesso ad un livello della rete, esistono numerosi tool (Wireshark, Cain and Abel, vomit, ecc.) in grado di catturare lo stream RTP in transito. Alcuni di essi permettono inoltre di convertire tale stream in file audio WAV (vomit ad esempio). Dopo aver registrato con successo una registrazione, è possibile estrarre da essa gli eventuali toni premuti. Ad esempio un semplice tool come

DTMF Decoder, è in grado di tradurre i toni emessi dalla scheda audio nei corrispondenti tasti che sono stati premuti.

**Difesa** L'unico modo per assicurare la riservatessa di una conversazione VoIP è quella di criptare la conversazione (in altre parole, lo stream RTP). Per fare ciò ci sono diversi modi. Il primo sarebbe quello attraverso il livello di rete, ovvero utilizzando IPSec. Un altro modo sarebbe quello di criptare i dati multimediali a livello del trasporto, utilizzando ad esempio Secure Real-time Transport Protocol.

# Capitolo 7

# Intercettare il Traffico VoIP

## 7.1 Introduzione

Qualsiasi parte mancante o alterazione dello stream VoIP, potrebbe cambiare drasticamente il significato di una conversazione. Rimuovere, sostituire o ripetere parole del parlato di una conversazione, può inevitabilmente avere conseguenze in varie contesti sociali. Se un maleintenzionato è situato tra le due parti di una conversazione ed è capace di intercettare e modificare il traffico, esiste una grossa varietà di cose che può fare.

In questo capitolo parleremo delle tecniche maggiormente utilizzare per intercettare e modificare una conversazione.

## 7.2 Tecniche di dirottamento tradizionali - Man-in-the-Middle

Con un tradizionale attacco man-in-the-middle un attaccante riesce ad inserire se stesso tra le due parti comunicanti, al fine di ascoltare e/o alterare i dati in transito senza che loro ne siano a conoscenza. In un tipico scenario VoIP, con un attacco di questo tipo, un hacker potrebbe intraprendere una varietà di ulteriori attacchi:

- Ascolto della conversazione
- Causare un denial of service effettuando il blackholing[1] della con-

[1]Con blackholing si intende l'operazione attraverso la quale si riesce a bloccare tutto il traffico per una certa destinazione inviandolo in un «buco nero» dove viene scartato. Tuttavia, con questo sistema, si scarta anche il traffico legittimo.

versazione

- Alterare la conversazione cancellando delle parti
- Alterare la conversazione ripetendo delle parti
- Alterare la conversazione inserendo delle parti
- Redirigere il traffico di una parte mittente verso un'altra parte ricevente

## 7.3 ARP Poisoning

L'ARP poisoning (detto anche ARP spoofing) è una tecnica di hacking che consente ad un attaccante, in una switched lan, di concretizzare un attacco di tipo man in the middle verso tutte le macchine che si trovano nello stesso segmento di rete. L'ARP poisoning è oggi la principale tecnica di attacco alle lan commutate. Consiste nell'inviare intenzionalmente e in modo forzato risposte ARP contenenti dati inesatti o, meglio, non corrispondenti a quelli reali. In questo modo la tabella ARP (ARP entry cache) di un host conterrà dati alterati (da qui i termini poisoning, letteralmente avvelenamento e spoofing, raggiro). Molto spesso lo scopo di questo tipo di attacco è quello di redirigere, in una rete commutata, i pacchetti destinati ad un host verso un altro al fine di leggere il contenuto di questi per catturare le password che in alcuni protocolli viaggiano in chiaro.

**Funzionamento** Questo attacco si basa su una debolezza intrinseca nel protocollo ARP: la mancanza di un meccanismo di autenticazione.

Ethernet, il più diffuso standard per le reti locali, identifica gli host in base ad un indirizzo a 48 bit chiamato MAC a differenza di Internet dove ciascun host viene mappato grazie ai 32 bit del protocollo IP.

Il protocollo ARP si occupa di gestire l'associazione tra indirizzi IP e indirizzi MAC. Quest'associazione, in Ethernet, viene fatta prima di ogni tipo di comunicazione. Sono previsti due tipi di messaggi dal protocollo ARP: ARP request (effettuata in broadcast) e ARP reply (effettuata in unicast). Un ipotetico host 192.168.1.1 che vuole comunicare con l'host 192.168.1.2 manderà una ARP request in broadcast con il proprio MAC il proprio indirizzo IP e l'indirizzo IP di destinazione; quando 192.168.1.2 riceverà l'ARP request risponderà con un'ARP reply destinato al MAC sorgente e contenente il proprio MAC. Per ottimizzare le prestazioni e

limitare il traffico queste informazioni vengono memorizzate nella tabella ARP (ARP cache) di ciascun host, così che non sia necessario effettuare continue richieste. Per migliorare ancora di più le prestazioni quando si ricevono delle ARP reply (alcuni anche con le ARP request), anche se non sollecitate, gli host aggiornano le informazioni della propria ARP cache.

Ora si analizzi il seguente scenario:

- Attaccante: IP = 192.168.1.2, MAC = 00:00:00:ZZ:ZZ:ZZ
- John: IP = 192.168.1.13, MAC = 00:00:00:JJ:JJ:JJ
- Linus: IP = 192.168.1.88, MAC = 00:00:00:LL:LL:LL

Le ARP cache di ciascun host prima dell'attacco saranno:

- Per l'attaccante

  192.168.1.2, MAC = 00:00:00:ZZ:ZZ:ZZ

  192.168.1.13, MAC = 00:00:00:JJ:JJ:JJ

  192.168.1.88, MAC = 00:00:00:LL:LL:LL

- Per John

  192.168.1.2, MAC = 00:00:00:ZZ:ZZ:ZZ

  192.168.1.13, MAC = 00:00:00:JJ:JJ:JJ

  192.168.1.88, MAC = 00:00:00:LL:LL:LL

- Per Linus

  192.168.1.2, MAC = 00:00:00:ZZ:ZZ:ZZ

  192.168.1.13, MAC = 00:00:00:JJ:JJ:JJ

  192.168.1.88, MAC = 00:00:00:LL:LL:LL

Per realizzare l'ARP poisoning l'attacker invierà delle ARP reply appositamente fatte: a John invierà una reply che ha come IP quello di Linus (192.168.1.88) ma come MAC il proprio (00:00:00:ZZ:ZZ:ZZ), a Linus invierà una reply con IP quello di John (192.168.1.13) e con MAC, anche questa volta, il proprio (00:00:00:ZZ:ZZ:ZZ). Per protrarre l'attacco è necessario inviare delle ARP reply ogni 10 secondi poiché spesso i sistemi operativi cancellano sistematicamente le voci dell'ARP cache.

Quindi dopo l'attacco le ARP cache di ciascun host saranno:

- Per l'attacker

  192.168.1.2, MAC = 00:00:00:ZZ:ZZ:ZZ

  192.168.1.13, MAC = 00:00:00:JJ:JJ:JJ

  192.168.1.88, MAC = 00:00:00:LL:LL:LL

- Per John

  192.168.1.2, MAC = 00:00:00:ZZ:ZZ:ZZ

  192.168.1.13, MAC = 00:00:00:JJ:JJ:JJ

  192.168.1.88, MAC = 00:00:00:ZZ:ZZ:ZZ

- Per Linus

  192.168.1.2, MAC = 00:00:00:ZZ:ZZ:ZZ

  192.168.1.13, MAC = 00:00:00:ZZ:ZZ:ZZ

  192.168.1.88, MAC = 00:00:00:LL:LL:LL

Le due vittime John e Linus crederanno di comunicare tra di loro, ma in realtà comunicheranno con l'attacker il quale inoltrerà il traffico proveniente da John verso Linus e viceversa il traffico proveniente da Linus verso John, realizzando così un MITM.

Dopo aver concretizzato il MITM, l'attacker sarà in grado di modificare, leggere e creare nuovi pacchetti. Potrà quindi facilmente sniffare tutto il traffico in chiaro come password telnet, ftp, pop3, irc, ecc.

### 7.3.1 Contromisure

L'utilizzo di IPv6, IPsec o di tabelle ARP statiche sono metodi che possono rivelarsi una difesa efficace contro attacchi di tipo ARP spoofing. Ovviamente è impensabile mantenere aggiornate le tabelle ARP di ogni host in una rete di grande dimensioni. Altre soluzioni potrebbero essere:

- usare un software come arpwatch che esamina le attività di rete e ne evidenzia le discordanze o come OpenAAPD, un demone anti ARP poisoning per OpenBSD o ancora un intrusion detection system (IDS) come Snort.

- usare il port security sugli switch ovvero fare in modo che per ciascuna porta del dispositivo possa esserci solo un MAC address.

- SARP ovvero Secure ARP, un'estensione del protocollo ARP che si basa sulla crittografia asimmetrica, così da poter autenticare il mittente.
- implementare 802.1x nella propria rete.

Avendone la possibilità (è necessario avere switch che supportano 802.1x e un server RADIUS), di sicuro la migliore soluzione è implementare 802.1x.

# Capitolo 8

# Autenticazione

## 8.1 Introduzione

I requisiti di autenticazione e autorizzazione sia per utenti che dispositivi, crescono di pari passo all'evoluzione VoIP. Gli utenti spesso mantengono identità multiple. Impiegano username e password differenti per vari contesti online, e spesso ciò causa problemi nel ricordarli. Per tali motivi VoIP ed altri servizi di rete contemporanei hanno la necessità di una gestione delle identità e la base di tale gestione sono i servizi di autenticazione.

Le identità di utenti e dispositivi non coincidono e per tale motivo hanno la necessità di essere verificate indipendetemente. Dal momento specifica SIP non definisce esplicitamente dei meccanismi di autenticazione, gli sviluppatori hanno adottato dei meccaniscmi di autenticazione utilizzati già in altre applicazioni. Il protocollo di autenticazione che gli amministratori di una rete VoIP maggiormete utilizzano è 802.1x/EAP.

802.1x e 802.11i/WPA2 si basano su un server di autenticazione (di solito un server RADIUS) e su un Autenticatore (uno switch o un access point wireless). 802.1x fornisce supporto per EAP(Extensible Authentication Protocol), il quale fornisce un framework per metodi di autenticazione multipli, che includono password, Kerberos, certificati digitali e autenticazione a chiave pubblica.

In questo capitolo illustreremo i principali protocolli utilizzati per autenticare le parti coinvolte in una comunicazione VoIP.

## 8.2 802.1x

Il protocollo 802.1x è usato per fornire un accesso autenticato alla rete. Anche se questo standard è stato progettato per le reti Ethernet, è stato adattato per essere usato su reti 802.11. In sostanza è uno standard per utilizzare EAP in una LAN wired o wireless.

802.1x vieta a client non autorizzati di connettersi alla LAN. Il client si deve prima autenticare con un server di Autenticazione (di solito un server RADIUS), prima che la porta sullo switch possa essere resa disponibile e la rete acceduta. EAP (Extensible Authentication Protocol) è un protocollo generale di autenticazione che fornisce un framework per meccanismi di autenticazione multipli come le tradizionali password, Kerberos, certificati digitali e autenticazioni a chiave pubblica.

### 8.2.1 Autenticazione 802.1x/EAP

Prima di spiegare il meccanismo di autenticazione 802.1x/EAP definiamo alcuni termini chiave.

**Supplicant** Con questo termine ci riferiamo al client in uno scambio EAP, ovvero il nodo che deve essere autenticato dall'autenticatore.

**Autenticatore** L'autenticatore è un access point wireless (AP) oppure uno switch (NAS - Network Access Server). L'autenticatore mantiene la rete chiusa a tutto il traffico non autenticato.

**Server di Autenticazione** Il server di autenticazione esegue l'autenticazione vera e propria del client e comunica all'autenticatore di consentire o negare il traffico a tale client. Il server di autenticazione è di solito un server RADIUS.

La figura 8.1 mostra il flusso di messaggio in uno schema di autenticazione 802.1x/EAP. Per prima cosa il client (una workstation, access point wireless, telefono IP, ecc.) invia una o più richieste al NAS. Il NAS (passo 2) inoltra i messaggi EAP al server di Autenticazione. Nel terzo passo, il server di autenticazione richiede le credenziali del supplicant (client) specificando il tipo di credenziali necessarie (bisogna notare che nella figura la freccia tra il server di autenticazione e il client, rappresenta una connessione logica e non fisica. Tutto il traffico passa attraverso il NAS). Nel quarto passo, il supplicant invia le sue credenziali al server di Autenticazione. Nel quinto passo, il server di autenticazione inoltra

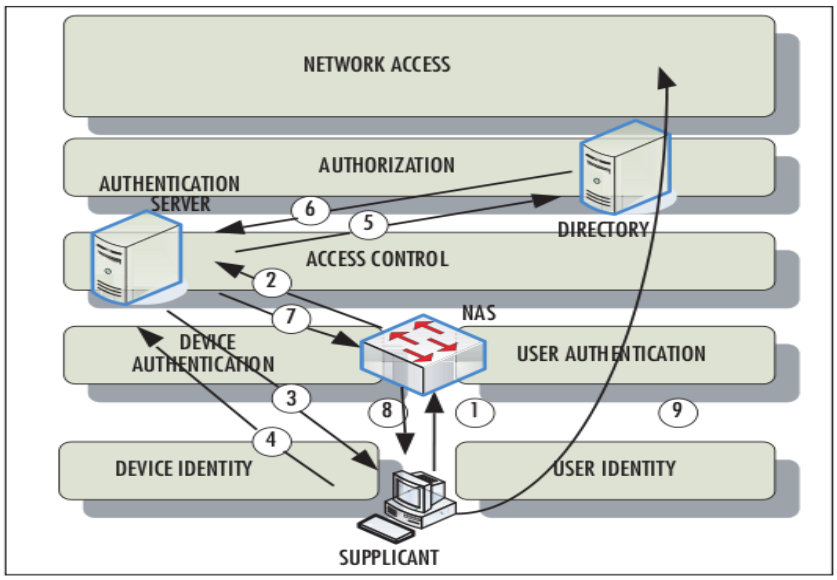


Figura 8.1: Autenticazione EAP con Autorizzazione

la richiesta di accesso ad un server di Directory, il quale risponde con un messaggio si successo o fallimento. In caso di successo, il server di directory comunica al server di autenticazione il dominio di appartenenza delle credenziali del client esaminato.Una volta aver validato le credenziali del supplicant, il server di Autenticazione trasmette una messaggio di successo o di fallimento al NAS (passo 7). Al passo 8, se l'accesso è concesso, il NAS apre la porta a tutto il traffico e lo scambio di dati tra il dispositivo LAN autenticato e la rete LAN viene permesso. Se l'accesso è consentito, allora (passo 9) il supplicant ha la facoltà di accedere alle risorse di rete.

In questo modo l'amministratore della rete può limitare l'accesso ad un client solamente ad una specifica VLAN. Le specifiche dell'autenticazione e dell'autorizzazione dipende dal tipo di politica EAP scelta.

## 8.3 Tipi di Autenticazione EAP

I diversi tipi di autenticazione EAP sono costituiti da due componenti: un tipo di autenticazione esterna e uno interno. Il tipo esterno definisce il metodo usato per stabilire un canale criptato tra il client (peer) e il server di Autenticazione.

Nella figura 8.2 viene usato un metodo di autenticazione esterno,

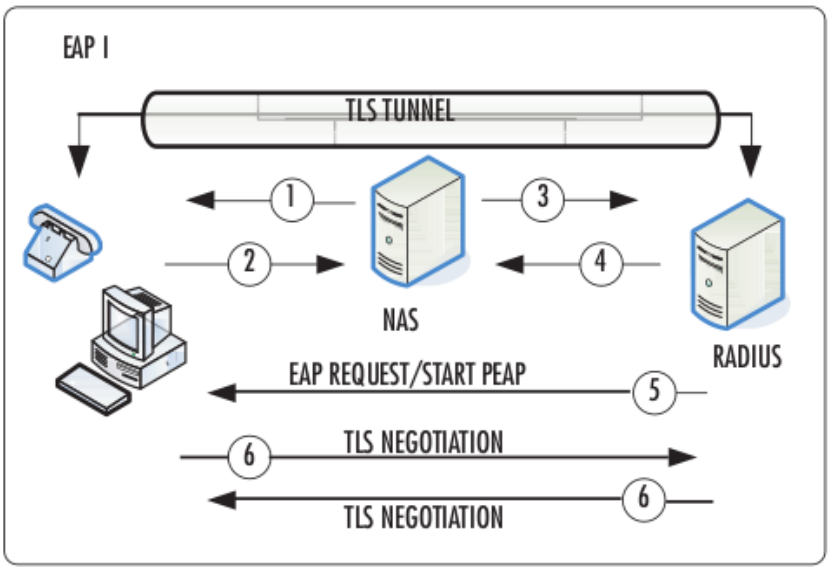


Figura 8.2: EAP Parte 1

PEAP, tra il client che può essere un telefono IP o una workstation, e il server di autenticazione RADIUS. Il NAS fa da intermediario ai primi scambi e ha la funzione di mediare passivamente il traffico in entrambe le direzioni. Il tunnel esterno verifica il server al client usando certificati digitali. Una volta che il canale esterno è stabilito, l'autenticazione interna passa le credenziali dell'utente al server di Autenticazione attraverso il tunnel TLS (Figura 8.3). Inviare le credenziali dell'utente attraverso il tunnel TLS protegge i dati dalla loro esposizione.

Una delle potenziali vulnerabilità di sicurezza di EAP, deriva dal fatto che i dati scambiati all'interno di qualche autenticazione esterna, sono scambiati in chiaro. Questo potrebbe portare in un attacco di tipo DoS, in quanto un attaccante potrebbe inondare la connessione con differenti tipi di messaggi di notifica EAP.

### 8.3.1 EAP-MD5

MD5 è l'equivalente del CHAP in cui un algoritmo hash a senso unico è utilizzato in combinazione con un segreto condiviso e una richiesta di identificazione per verificare che il richiedente è a conoscenza del segreto condiviso. MD5 è considerato un metodo di autenticazione di livello base e generalmente non appropriato in caso sia necessario un alto livello di sicurezza per la protezione di beni di grande valore. Questo accade per

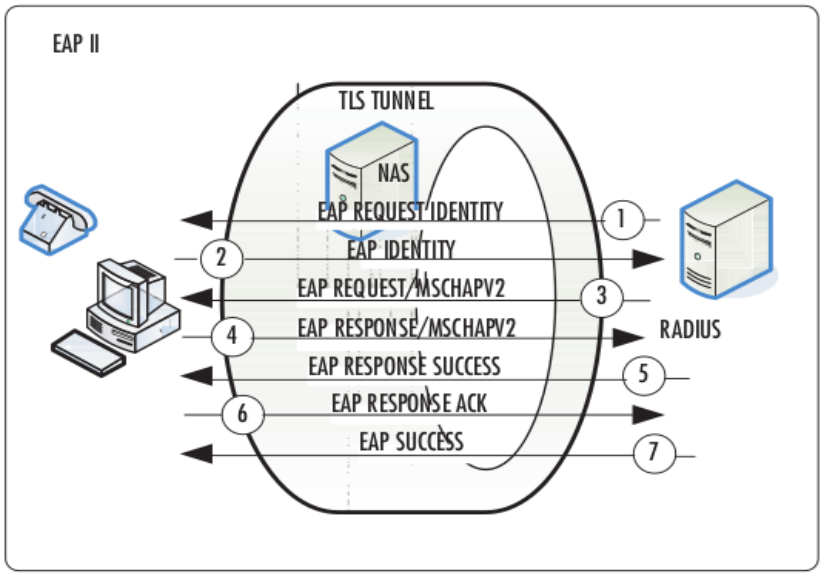


Figura 8.3: EAP Parte 2

diverse ragioni. Come ogni metodo che utilizza richieste random e un algoritmo hash, è vulnerabile agli attacchi basati su dizionario. Se un attaccante riesce ad ottenere la richiesta e la risposta hash, è in seguito possibile eseguire un programma off-line con lo stesso algoritmo del richiedente, inserendo parole contenute in un dizionario fino a quando la risposta hash coincide con quella del richiedente. A questo punto l'attaccante conoscerà la password del richiedente e potrà sottrarne l'identità per ottenere l'accesso alla rete. Questo procedimento risulta ancora più semplice nelle wireless LAN, dove la richiesta e la risposta viaggiano nell'aria. Questo è il motivo per cui è importante scegliere password che non siano parole di senso compiuto. In aggiunta, EAP-MD5 offre soltanto l'autenticazione lato client (ovvero, il client viene autenticato alla rete). Altri metodi EAP offrono mutua autenticazione per cui il client è autenticato alla rete e la rete è autenticata al client.

### 8.3.2 EAP-TLS

Il Transport Layer Security (TLS) offre un processo di autenticazione particolarmente sicuro, che sostituisce le semplici password con certificati lato client e lato server tramite l'utilizzo della infrastruttura a chiave pubblica (Public Key Infrastructure o PKI). È supportata la mutua autenticazione, e le chiavi di sessione dinamiche. TLS è una buona scelta

quando si richiede un elevato livello di autenticazione e sicurezza ed è presente una infrastruttura a chiave pubblica. Comunque, l'utilizzo di una PKI, in cui ciascun client ha il suo proprio certificato, è oneroso se comparato ai sistemi basati su password. Tale onere deriva dagli strumenti software richiesti affinché il sistema sia efficace.

### 8.3.3 EAP-TTLS

Tunnelled Transport Layer Security (TTLS) è un'estensione del TLS ed è stato sviluppato per superare la necessità, generata dal TLS, di certificati lato client (sono invece richiesti certificati lato server). Così come l'altro dei due metodi attualmente disponibili di autenticazione tramite tunnel (l'altro è il PEAP), TTLS è un metodo a due passaggi. Nel primo, un algoritmo asimmetrico basato sulle chiavi del server è utilizzato per verificare l'identità del server e per creare il tunnel di crittazione simmetrica. Il secondo passaggio riguarda la verifica dell'identità del client utilizzando un secondo metodo di autenticazione tramite il tunnel di crittazione simmetrica per l'attuale negoziazione dell'autenticazione. Questo secondo metodo di autenticazione utilizzato con il tunnel può essere un tipo di EAP (spesso MD5) o un metodo di vecchio tipo come PAP, CHAP, MS-CHAP, o MS-CHAP V2. Il tunnel a crittazione simmetrica del TTLS è utilizzato solo per proteggere il metodo di autenticazione del client. Una volta verificato, il tunnel collassa.

### 8.3.4 EAP-LEAP

Lightweight Extensible Authentication Protocol sviluppato dalla Cisco, deriva da EAP. LEAP si basa su un protocollo di autenticazione chiamato reciproco consenso che sta a dire in poche parole che sia il client sia l'access point a cui il client richiede la connessione devono autenticarsi prima di avere accesso all'interno della rete. In questo modo si previene l'accesso non autorizzato di access point estranei alla rete.

## 8.4 Metodi di Autenticazione Interni

Esistono anche diversi metodi di autenticazione interni, di cui i pù utilizzati sono quelli che derivano da CHAP.

### 8.4.1 CHAP

CHAP (Challenge-Handshake Authentication Protocol) è un protocollo di autenticazione che identifica un utente presso un Internet Service Provider.

CHAP è uno schema d'autenticazione usato dai server PPP per convalidare l'identità dei client remoti. CHAP verifica periodicamente l'identità del client tramite un processo handshake. Ciò accade non appena viene stabilito il primo contatto e può accadere di nuovo in qualunque momento. La verifica si basa su un segreto condiviso (come la password dell'utente).

1. Dopo aver stabilito una connessione, il client invia il proprio identificativo utente ed il server risponde con una domanda di sfida (challenge), costituita da un numero pseudocasuale. 2. Il client esegue l'hash (può essere un MD5) del challenge assieme alla sua password e lo reinvia. 3. Il server, che conosce la password, è in grado di eseguire lo stesso calcolo e quindi comparare i due valori verificando la correttezza del valore ricevuto. Se i valori non combaciano la connessione viene terminata. 4. Ad intervalli casuali il server ripropone un challenge al client e vengono ripetuti i primi tre passi.

In questo modo la chiave non circola sulla rete, ma rimane il problema della conoscenza della chiave che deve essere nota per entrambi i sistemi. Il CHAP offre però una protezione contro gli attacchi replay grazie all'uso di un identificatore e di un challenge variabili. Questo protocollo necessita che il client conservi la password in formato testo.

### 8.4.2 MS-CHAP

MS-CHAP è un protocollo sviluppato da Microsoft deriva direttamente dal CHAP ed è stato modificato per integrare la funzionalità di autenticazione basate sulla cifratura, tra Pc in reti windows.

MS-CHAP eredita dal CHAP il meccanismo di challenge-response per autenticare attraverso una password il client verso il server. Il server di autenticazione diversamente dal CHAP non ha necessità di memorizzare la password neanche in formato cifrato, e gli algoritmi utilizzati per il processo di cifratura e per quello di hashing sono rispettivamente il DES (Data Encryption Standard) e l'MD4 (Message Digest 4).

### 8.4.3 MS-CHAPv2

MS-CHAP versione 2 deriva dal protocollo MS-CHAP. In MS-CHAPV2 il processo di autenticazione è reciproco, il client e il server si presentano e il server deve dimostrare al client che è in grado di accedere al database dove è contenuta la password dell'utente che sta tentando la connessione.

In linea di massima i passi compiuti dal processo di autenticazione sono:

1. Il client contatta il server e stabilisce una sessione. 2. Il server di autenticazione invia al client un messaggio composto da: 1. un identificatore della sessione(IdS) 2. una stringa pseudocasuale (A). 3. Il client riceve il messaggio dal server e invia una risposta composta da: 1. Username 2. Una stringa fittizia(B). 3. La stringa pseudocasuale(A), l'identificativo di sessione (IdS), la password utente tutto cifrato. 4. Il server riceve il messaggio dal client lo verifica e invia la relativa risposta composta da: 1. L'esito del tentativo di connessione. 2. La stringa fittizia(B) e la password utente tutto cifrato. 5. Il client riceve la risposta e se è avvenuta l'autenticazione utilizza la sessione altrimenti interrompe la connessione.

Rispetto al MS-CHAP l'MS-CHAP v2 è più sicuro questo perché:

1. Ogni volta che l'utente si collega è generata una chiave per crittografare i dati basata sia sulla password utente sia su una stringa casuale. Nella versione v1 essendo la chiave generata solo a partire dalla password la chiave di crittografia è sempre la stessa. 2. I dati inviati e quelli trasmessi sono cifrati con chiavi generate separatamente e non come nella versione v1 dove la chiave era sempre la stessa.

# Capitolo 9

# Separare logicamente il traffico di Rete

## 9.1 Introduzione

Uno dei principali vantaggi nell'unire voce e dati è quello di risparmiare denaro e di semplificare l'amministrazione e la gestione della rete, in quanto entrambi i tipi di traffico viaggiano sulla stessa infrastruttura. Tuttavia molti ingegneri, durante la fase di progettazione di una rete VoIP, fanno numerosi sforzi nel cercare di dividere logicamente il traffico voce da quello dati.

Al livello 2 e 3 del modello ISO/OSI, i pacchetti voce sono pressochè indistinguibili da quelli dati, e per questo soggetti agli stessi rischi di sicurezza delle reti solo a traffico dati. L'idea di fondo alla base della separazione logica del traffico risiede nel fatto che eventi tipo congestione, fenomeni di sicurezza come worm, attacchi DoS che infettano una rete, non abbiano effetto sull'altra.

La figura 9.1 mostra una rete con traffico VoIP e dati, illustrando le principali componenti di sicurezza coinvolte nella separazione logica dei due tipi di traffico. Tra l'esterno e l'interno della rete è presente un primo livello di sicurezza mediante l'utilizzo di firewall e router con ACL (Access Control List). All'interno della rete, il traffico VoIP viene diviso dal traffico dati utilizzando VLAN, indirizzi IP privati, NAT e firewall VoIP-aware.

In questo capitolo daremo una panoramica sulle Virtual LAN, discutendo di come il loro utilizzo possa aumentare la sicurezza di un'infrastruttura VoIP

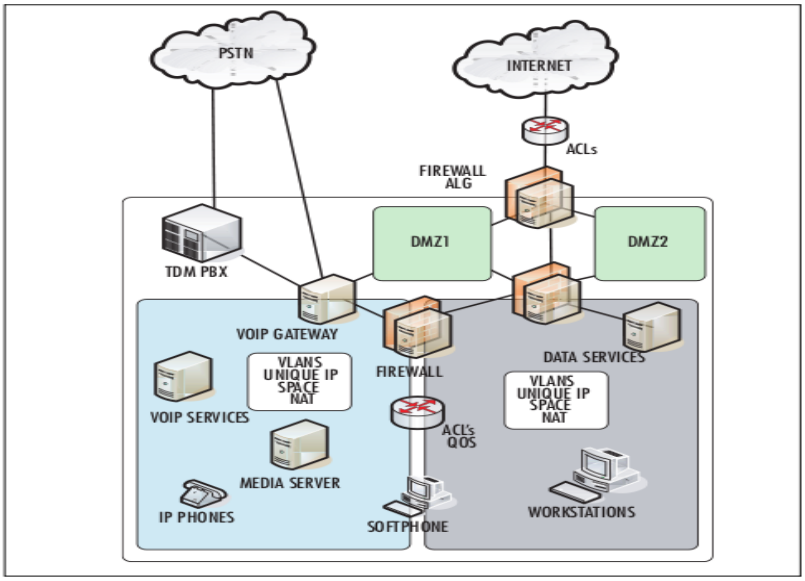


Figura 9.1: Rete di riferimento VoIP

## 9.2 VLAN

Una VLAN é una rete di computer che si comportano come se fossero connessi allo stesso cavo, malgrado essi siano connessi a diversi segmenti di LAN (quindi, su differenti Domini di Collisione). Il Network Administrator può configurare VLAN sia tramite software, sia tramite hardware, che le rende estremamente flessibili. Uno dei più grossi vantaggi delle VLAN emerge quando un computer viene fisicamente cambiato di locazione: esso rimane comunque collegato alla stessa VLAN senza alcuna riconfigurazione dell'hardware.

Il mondo delle VLAN é attualmente dominato dall'IEEE 802.1Q tagging protocol. Prima di esso esistevano già altri protocolli proprietari, come l'ISL di Cisco (Inter-Switch Link, una variante dell'IEEE 802.10) e il VLT di 3Com (Virtual LAN Trunk). Attualmente si tende ad abbandonare i protocolli proprietari in favore dell'802.1Q.

Inizialmente i progettisti di rete configuravano le VLAN con lo scopo di ridurre le dimensioni del Dominio di Collisione in un ampio segmento Ethernet, aumentando di conseguenza le performance globali. Quando però gli Switch fecerò scomparire questo problema poiché, in pratica, il Dominio di Collisione era simulato e non più reale, l'attenzione fu rivolta

a ridurre le dimenzioni del Dominio di Broadcast al livello MAC.

Le VLAN possono essere utili anche allo scopo di restringere l'accesso a delle risorse, senza bisogno di modificare la topologia fisica della rete.

Le VLAN operano al livello 2 (il Data Link Layer) dello Stack ISO/OSI2. E' possibile però configurare delle VLAN costruite mappando direttamente gli indirizzi IP, o delle sottoreti intere, coinvolgendo, di fatto, anche il livello 3 (il Network Layer).

Nel contesto delle VLAN, il termine "trunk3" denota un collegamento della rete che trasposta VLAN multiple, identificate tramite etichette (dette "tag") inserite nei loro pacchetti. Ogni "trunk" deve passare attraverso le "tagged-port" di una device abilitata alle VLAN: spesso si tratta collegamenti Switch-Switch o Switch-Router.

In pratica una VLAN viene realizzata simulando un unico Dominio di Broadcast: data una VLAN (che chiameremo VLAN A), i pacchetti che partono da macchine che appartengono ad A sono diretti a tutte e sole le macchine di A.
Il suo funzionamento é semplice e, al contempo, potente.

**Possibili configurazioni** Gli amministratori di rete possono configurare VLAN in vari modi:

- a livello protocollo, usando IP, IPX, ecc.: lo Switch analizza il frame di livello 2 ed il relativo campo "protocol" e dirige il traffico verso la rispettiva VLAN - coinvolti i livelli 2/3;
- basandosi sul MAC address delle macchine: lo Switch é configurato con tabelle che raggruppano i MAC address in VLAN e dirige il traffico in base ad esse - coinvolto il livello 2;
- basandosi sulle subnet IP: simile alla tecnica basata su MAC, tranne appunto che si usa l'indirizzo IP - coinvolto il livello 3;
- basandosi sulle porte degli Switch che devono gestire la VLAN - coinvolto il livello 1;

**Metodi di identificazione su VLAN** Quando uno Switch é configurato per supportare più VLAN basate su livelli superiori al primo, esiste

la necessità di identificare ogni singolo pacchetto (di livello 2 o 3, a seconda delle esigenze) per poterlo "dirigere" da e verso la propria VLAN. Per astrazione si possono indicare 2 metodologie per identificare una VLAN: il Frame-Tagging e il Frame-Filtering:

1. il Frame-Tagging si basa sulla modifica delle informazioni del frame di livello 2, così che lo Switch possa dirigere il traffico verso la VLAN corretta, dopo aver riportato il frame in condizioni normali - necessità di modifica dei frame;

2. il Frame-Filtering fa si che lo Switch analizzi i pacchetti di livello 2 in base ad un particolare criterio e diriga il traffico di conseguenza - nessuna necessità di modifica dei frame;

### 9.2.1 Sicurezza delle VLAN

La sicurezza delle VLAN e del livello 2 è un argomento complesso. L'unica regola generale che bisogna obbligatoriamente rispettare è che individui non autorizzati non abbiano accesso alla console dello switch. Per un accesso da terminale invece bisogna prevedere meccanismi di autenticazione robusti (RADIUS).

Il funzionamento delle VLAN dipende dalla presenza o meno delle informazioni nel tag. Per riuscire a garantire l'integrità di tali informazioni, l'idea è quella di certificarle.

### 9.2.2 VLAN e Softphone

In un ambiente VoIP, i softphone presentano delle sfide di sicurezza, ed in particolar modo se le VLAN sono impiegate come maggiore controllo di sicurezza. Diversi softphone popolari (come X-Lite) memorizzano le credenziali all'interno del registro di Windows in chiaro, anche dopo la disinstallazione del programma. Altri softphone invece, contengono al loro interno software pubblicitario che raccoglie informazioni private dell'utente. Firewall o IDS in questo caso hanno un uso limitato in quanto i softphone richiedono che sui firewall ci siano un gran numero di porte UDP aperto.

La regola più importante per rendere sicuri i softphone è quella di rafforzare il sistema operativo sottostante. Qualsiasi malware che infetta una qualsiasi applicazione del sistema, potrebbe potenzialmente interferire anche nelle comunicazioni vocali.

# Capitolo 10

# Crittografia in Voip

## 10.1 Introduzione

Ci sono due protocolli di segnalazione VoIP concorrenti, H.323 definito dall' ITU e SIP definito dall' IETF. Di conseguenza, ci sono anche due gruppi di protocolli per garantire la sicurezza. Il primo per H.323 è un gruppo di protocolli denominato H.235.x, mentre l'altro per SIP include TLS, S/MIME e SRTP. Tali gruppi non sono totalmente esclusivi, nel senso che alcuni protocolli(TLS, SRTP) sono presenti in entrambe le suite. In questo capitolo, concentreremo la nostra attenzione sul gruppo di protocolli per SIP definiti dall'IETF.

## 10.2 Suite di protocolli dell'IETF

Per rendere sicuro il sistema VoIP, l'IETF(Internet Engineering Task Force) ha incluso nella specifica SIP tre protocolli di sicurezza già noti: Transport Layer Security (TLS), Secure/Multiporpose Internet Mail Extensions (S/MIME) e Secure Real-Time Transfer Protocol (SRTP). L'approccio adottato è stato quello di includere un nuovo layer di sicurezza sottostante la specifica del protocollo VoIP esistente, anzichè riscrivere per intero un nuovo protocollo VoIP sicuro. Il vantaggio di questo approccio è stato quello che le implementazioni esistenti, potevano essere riusate per comunicazioni sicure semplicemente aggiungendo il nuovo layer di sicurezza. In generale, TLS, che è stato scelto per proteggere i messaggi di segnalazione SIP, offre un livello di sicurezza superiore alla sua entità peer. É fondamentalmente un successore di SSL (Secure Sockets Layer) versione 3. Il Service Data Unit (SDU) viene cifrato dal livello superiore prima della trasmissione. All'altra estremità, il Protocol

Data Unit ricevuto (PDU) è decifrato e passato al livello superiore. Ciascuna entità a entrambe le estremità deve avere un certificato autentico rilasciato da una Certificate Authority (CA), obbligatorio per il funzionamento dell' operazione di handshake di TLS. I messaggi di segnalazione SIP, vengono passati attraverso i tutnnel sicuri TLS.

SRTP è utilizzato per garantire la sicurezza del traffico voce/video da eventuali manomissioni e ascolti indiscreti. Esso garantisce la riservatezza dei payload RTP e l'integrità di tutti i pacchetti RTP mediante l'adozione, come impostazione predefinita, di Advanced Encryption Standard (AES) come algoritmo per la cifratura/decifratura utilizzando una chiave simmetrica. Inoltre protegge contro attacchi a pacchetto ripetuto. La questione più delicata dell'uso di SRTP è il meccanismo mediante il quale la chiave segreta possa essere condivisa tra due nodi della comunicazione. Incorporare la chiave manualmente in tutti i telefoni cellulari sarebbe troppo pesante e suscettibile a errori. Per ragioni di efficienza, RTP e SRTP possono essere implementati come un unico strato, piuttosto che due distinti livelli. TLS e SRTP sono i componenti chiave che giocano un ruolo importante nel garantire la sicurezza del servizio VoIP.

Tuttavia, ci devono essere protocolli di supporto oppure un'infrastruttura in grado di autenticare gli utenti, convalidare nodi e certificati degli utenti, e scambiare chiavi crittografiche. Ciascuno di questi elementi dovrebbero lavorare insieme in armonia per fornire servizi VoIP sicuri.

## 10.3 S/MIME: Autenticazione dei Messaggi

Per poter garantire la sicurezza della posta elettronica, il messaggio deve essere protetto da eventuali manomissioni, e sia il mittente che il destinatario devono anche essere correttamente identificati. Il motivo per cui lo Spam è fiorente in questi giorni è che l'indirizzo e-mail del mittente può essere facilmente fraudolento o falso. Secure/Multipurpose Internet Mail Extensions (S/MIME), indicato nelle RFC 3850 e 3851,fornisce uno standard per la crittografia a chiave pubblica e per la firma di e-mail incapsulate nel popolare formato MIME.

S/MIME fornisce i seguenti servizi di sicurezza crittografica per le applicazioni di messaggistica elettronica: autenticazione, integrità del messaggio, non repudiazione(utilizzando firme digitali), e la riservatezza

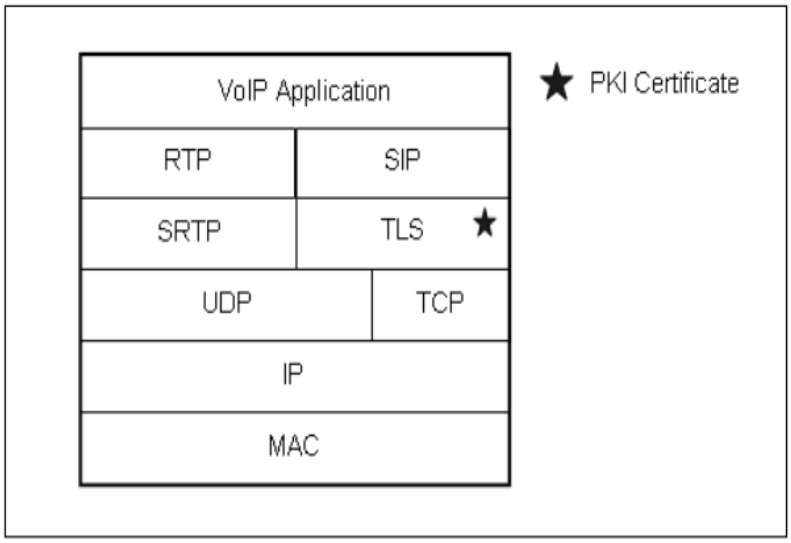


Figura 10.1: Livelli di Sicurezza VoIP

dei dati (utilizzando la crittografia). S/MIME non è limitato alla posta. Può essere utilizzato con qualsiasi meccanismo di trasporto che trasporti dati MIME, come ad esempio i body di messaggi HTTP o SIP.

Anche se S/MIME si applica al body del messaggio globale, lo standard SIP prevede un meccanismo per applicare S/MIME per proteggere anche gli header. Informazioni del body come SDP sono criptati con S/MIME per mantenere l'integrità e rimanere riservate. Tuttavia, informazioni nell'intestazione, come ad esempio A, Da, Call-ID, CSeq, non possono rimanere riservate end-to-end. Essi sono informazioni indispensabili per gli intermediari, come ad esempio SIP proxy server e firewall, per stabilire la chiamata richiesta. Per superare questo problema, le informazioni sono fornite sia in chiaro che in formato codificato S/MIME all'interno di un messaggio SIP. Così gli intermediari hanno accesso alle informazioni senza preoccuparsi di doverli prima decifrare, mentre il destinatario finale, con una chiave corretta, è in grado di decifrare le informazioni cifrate e di confrontarle con quelle in chiaro nel messaggio al fine di verificare l'integrità e l'identità del mittente.

Per capire come un sistema basato su S/MIME consegni messaggi

sicuri alla propria destinazione, è necessario la conoscenza dei meccanismi di base di un sistema di messaggistica PKI. La figura 10.2 mostra il funzionamento generale e il flusso di chiavi e messaggi sia all'interno del sistema che tra sistemi stessi.

1. Viene effettuato l'hash del messaggio con un algoritmo di hashing.
2. Alice firma l'hash del messaggio usando un algoritmo di firma digitale, ed aggiunge la firma al messaggio originale e al suo certificato
3. Viene generata casualmente una chiave di sessione per crittografare il messaggio, il certificato e la firma usando un algoritmo di crittografia.
4. La chiave di sessione viene crittografata con la chiave pubblica di Bob usando un algoritmo di crittografia a chiave pubblica ed aggiunta all'interno del messaggio cifrato. Il messaggio risultante, viene trasmesso al ricevitore.
5. Quando Bob riceve il messaggio, recupera la chiave di sessione decifrandola con la propria chiave privata, usando lo stesso algoritmo utilizzato al passo 4.
6. Con la chiave di sessione recuperata, vengono decifrati messaggio, certificato e firma utilizzando lo stesso algoritmo del passo 3. In questo modo la riservatezza dei dati è garantita. A questo punto, Bob deve verificare se il messaggio è stato realmente firmato da Alice e che non sia stato manomesso durante la trasmissione.
7. Usando lo stesso algoritmo del passo 1, Bob fa l'hash del messaggio.
8. Bob verifica che la firma di Alice sia autentica. Se lo è, viene recuperata la chiave pubblica di Alice dal suo certificato.
9. Bob usa lo stesso algoritmo del passo 2 e firma l'hash del messaggio calcolato con la chiave pubblica di Alice.
10. La firma calcolata viene confrontata con quella ricevuta. Se sono diverse, il messaggio è stato alterato. La manomissione è avvenuta al di fuori della rete. Autenticazione, integrità del messaggio e non repudiazione vengono quindi verificate.

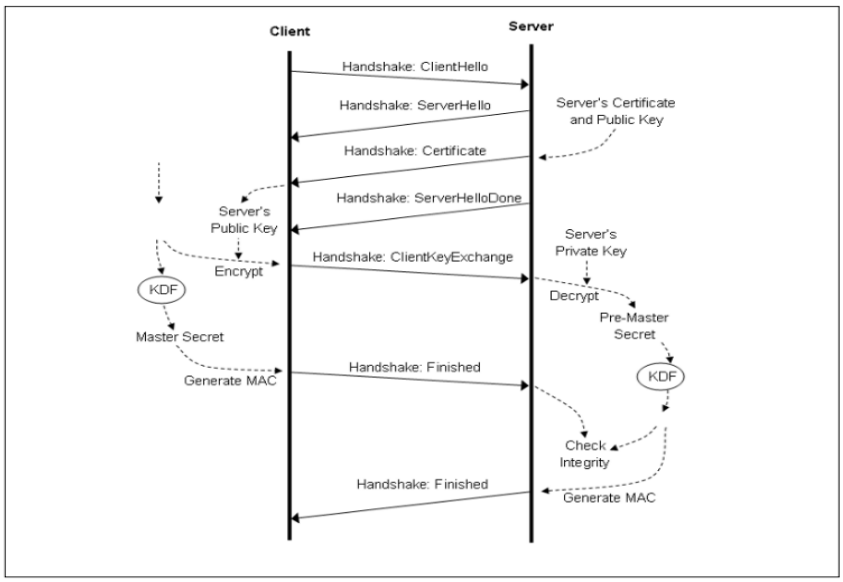


Figura 10.2: SSL HandShakes per i Certificati e lo Scambio delle Chiavi

## 10.4 TLS: Lo scambio delle Chiavi e la Sicurezza dei pacchetti di segnalazione

TLS è un protocollo che fornisce un canale sicuro tra due macchine. Ha delle caratteristiche per proteggere i dati in transito e per autenticare la propria controparte verificandone i certificati X.509. Tale canale sicuro è trasparente nel senso che i dati che passano su di esso non vengono modificati. I dati tra client e server sono crittografati, ma quello che una parte scrive è essattamente quello che l'altra parte legge. Come mostrato in figura 10.1, TLS è posizionato tra il livello TCP e il livello SIP, ovvero un messaggio al livello SIP è criptato usando TLS e trasmesso attraverso TCP. Ogni entità coinvolta in una comunicazione deve avere un certificato valido rilasciato da una CA.

### 10.4.1 Certificati e Scambio delle Chiavi

La figura 10.2 mostra il processo di handshake tra client e server. Lo scopo di tale processo è in primo luogo quello di far accordare client e server su un insieme di algoritmi che verranno usati per proteggere i dati. In secondo luogo, quello di stabilire un insieme di chiavi che verranno usate con tali algoritmi.

La figura 10.2 mostra il caso in cui il client sfida l'autenticazione del server:

1. Con i messaggi ClientHello e ServerHello, client e server si accordano su una lista di algoritmi che useranno.

2. Il certificato e la chiave pubblica del server sono contenuti nel messaggio Certificato.

3. Il client genera un numero casuale, chiamato Pre-Master Secret key. Prima di ricevere il messaggio Certificato, il client verifica l'autenticazione del certificato del server e ne estrae la chiave pubblica. La Pre-Master Secret key viene criptata con la chiave pubblica del server e inviata attraverso il messaggio ClientKeyExchange al server. Nel frattempo, la Key Derivation Function (KDF) genera una master key derivata dalla Pre-Secret Master key.

4. Il server decripta il messaggio ClientKeyExchange con la sua chiave privata, ottenendo la Pre-Master Secret Key. Usando la stessa Key Derivation Function del client, la master key viene derivata dalla Pre-Master Secret Key.

5. Con la master key, il client genera il Message Authentication Code (MAC) di tutti i messaggi precedentemente ricevuti dal server e lo invia al server nel messaggio Finished.

6. Con la master key, il server genera un MAC di tutti i messaggi precedentemente ricevuti dal client e lo invia al client in un messaggio Finished.

7. Sia il server che il client verificano l'integrità del MAC ricevuto confrontandolo con il MAC di tutti i messaggi che avevano precedentemente inviato.

8. Se ha verifica ha successo, sia il server che il client condividono la stessa Master Secret Key.

La figura 10.3 mostra come i dati dai livelli superiori vengono incapsulati dal livello TLS/SSL. Dopo essere stati frammentati, ai dati viene aggiunto il MAC e successivamente scriptati. Quindi il record header SSL/TLS, che contiene il content type, la lunghezza, la versione SSL, viene aggiunto al testo cifrato. Ci sono quattro diversi content type: application, alert, handshake e change cipher specification.

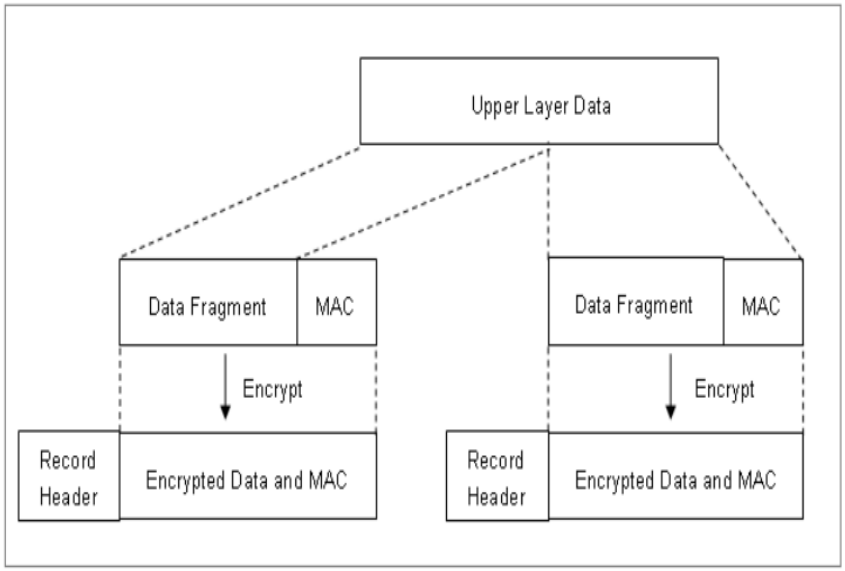


Figura 10.3: SSL/TLS

## 10.5 SRTP: Sicurezza dei Pacchetti Audio/-Video

SRTP, specificato nella RFC 3711, descrive come proteggere i canali di comunicazione audio/video, cifrando il payload del pacchetto RTP al fine di ottenere l'autenticazione dell'intero pacchetto e di proteggerlo da eventuali attacchi a replay:

1. La riservatezza dei pacchetti RTP garantisce che il payload venga letto soltanto da chi conosca la chiave di cifratura corretta.

2. L'autenticazione del pacchetto RTP protegge l'integrità dello stesso contro la falsificazione, l'alterazione o la sostituzione.

3. La protezione contro la ripetizione assicura che l'indirizzo della sessione (Indirizzo IP, porta UDP e Synchronization Source RC) non finiscano in un attacco DoS.

SRTP è situato tra l'applicazione RTP e il livello del trasporto RTP. Per garantire la riservatezza del payload RTP e l'integrità di tutti i pacchetti RTP viene utilizzato AES con delle chiavi di cifratura simmetriche. Il payload dall'applicazione RTP è criptato ed incapsulato in un pacchetto SRTP.

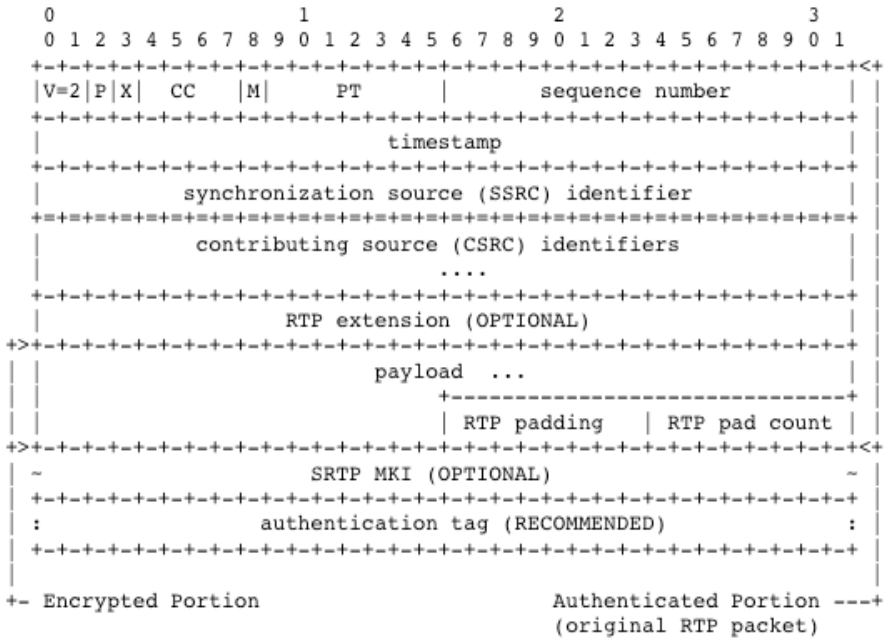


Figura 10.4: Formato del pacchetto SRTP

La problematica principale nell'usare SRTP è il modo in cui la chave segreta viene condivisa tra le due entità partecipanti alla comunicazione. La figura 10.4 mostra il formato del pacchetto SRTP.

La progettazione di SRTP ha tenuto in considerazione alcuni requisiti non funzionali (non relativi alla sicurezza) relativi a motivazioni di efficienza. Per raggiungere questi obiettivi, l'Internet Engineering Task Force (IETF) ha scelto di utilizzare la crittografia in modo da garantire che non si possa verificare nessuna espansione del payload durante questa fase. Inoltre, l'header TCP viene lasciato non cifrato, permettendo quindi ottimizzazioni trasmissive come ad esempio la compressione.

SRTP cifra soltanto il payload (audio o video) per garantire la riservatezza mentre l'algoritmo di autenticazione protegge l'integrità dell'intero pacchetto RTP originale. Il campo Master Key Identifier (opzionale) e il campo authentication tag(consigliato) sono gli unici campi che SRTP aggiunge al pacchetto RTP originale.

Il Master Key Identifier identifica quale master key è stata usata per derivare le chiavi di sessione correntemente in uso per la cifratura e/o per l 'autenticazione del pacchetto corrente. Anche se alcune volte non viene usato, il campo MKI è lungo 4 byte, ed è usato in un sistema che impiega scambi di chiavi multipli.

L'authentication tag ha una lunghezza variabile e garantisce che un attaccante non possa nè modificare pacchetti nello stream di dati nè inserirne di nuovi all'interno di esso. L'operazione di autenticazione è eseguita dopo l'operazione di cifratura e protegge l'intero pacchetto RTP. Dal momento che il numero di sequenza fa parte di questa protezione, l'authentication tag fornisce protezione anche contro gli attacchi a replay.

I metodi di cifratura Advanced Encryption Standard Counter Mode (AES-CM) e NULL sono obbligatori per implementare SRTP. Il metodo NULL è usato quando è richiesta soltanto l'autenticazione, tralasciando quindi la cifratura. Quando viene usato il metodo NULL, il payload originale RTP resta invariato. AES-CM invece è il metodo di cifratura predefinito utilizzato in SRTP. La ragione principale per la quale è stato scelto AES-CM è stata perchè non viene prodotta nessuna espansione del payload dopo che esso è stato cifrato. Un altra caratteristica di AES-CM è che esso permetto anche la cifratura di pacchetti non in ordine, il che implica la possibilità di processare pacchetti in parallelo.

HMAC-SHA1 (Hash-Based Message Authentication Code with Secure Hashing Algorithm 1) è l'algoritmo di autenticazione obbligatorio definito nella RFC 2104. Lo standard raccomanda che lo stream RTP debba essere protetto con un tag di autenticazione di 10 byte (80 bit). Una cosa da tenere in considerazione, è che payload comuni VoIP possono essere più piccoli di 20 byte, rendendo l'authentication tag il 50 percento dell'intera taglia del payload. Per ridurre questo overhead, può essere usato un authentication tag di 4 byte, ma soltanto se i rischi di sicurezza possono essere accettabili per la specifica applicazione.

SRTP può creare tutte le chiavi di autenticazione e di cifratura a partire da una singola master key. Per fare ciò, usa un algoritmo di derivazione delle chiavi basato su AES-CM. E' importate notare che SRTP non definisce lo scambio delle master key. SRTP non definisce alcun algoritmo di scambio delle chiavi. Ci sono numerose proposte a tale scopo, di cui l'implementazione più comune è quella del Security Description (SDES) protocol definito nella RFC 4568.

SRTP è un protocollo di sicurezza per internet conciso ed efficiente, che funziona bene ed ha raggiunto una buona interoperabilità. Tuttavia, la mancanza di un buon algoritmo per lo scambio delle chiavi ha impedito il diffondersi di una vasta varietà di implementazioni.

### 10.5.1 Multimedia Internet Keying

Multimedia Internet Keying (MIKEY) è una semplice soluzione per la gestione delle chiavi. Fornisce tre modi differenti per trasportare o stabilire le chiavi: con l'uso di chiavi precondivise, con la crittografia a chiave pubblica e con lo scambio di chaivi Diffie-Hellman.

### 10.5.2 Session Description Protocol Security Descriptions

SDP Security Description specifica un nuovo attributo SDP denominato *crypto*, usato per segnalare e negoziare i parametri crittografici dello stream SRTP. La definizione dell'attributo crypto è limitata al caso di stream 1-1 unicast. Si assume che il protocollo del trasporto sottostante sia sicuro (IPSec, TLS, SIP S/MIME) protegga il messaggio SDP che contiene l'attributo crypto. L'attributo descrive parametri crittografici, paramentri della chiave e parametri della sessione.

### 10.5.3 Riservatezza

La riservatezza è ottenuta cifrando il payload in modo tale che soltanto coloro in possesso della chiave possano leggerlo. Nella figura 10.5 $B_{i,j}$, è la cifratura AES del valore iniziale (IV) con la chiave.

Ogni IV è cifrato insieme con ogni chiave per produrre un blocco pseudocasuale a 128 bit mostrato come $B_{i,j}$. Ogni blocco di 128 bit viene messo in or esclusivo con il blocco del payload RTP associato per produrre un blocco di cifrato.

### 10.5.4 Autenticazione dei Messaggi

L'integrità dei messaggi è garantita dall'utilizzo di una funzione hash in congiunta con una chiave segreta. Come mostrato in figura 10.6, sul header e sul payload del pacchetto SRTP viene calcolata la funzione one-way Hash-Based Message Authentication Code with Secure Hashing Algorithm 1 (HMAC-SHA1) insieme ad una chiave segreta. Il mittente scrive l'hash HMAC-SHA1 nel tag autenticazione. Il ricevitore dopo aver computato lo stesso valore sul pacchetto ricevuto, ne verifica l'uguaglianza rispetto al valore contenuto nel tag autenticazione, accettando il pacchetto se i due valori sono uguali e rifiutandolo altrimenti.

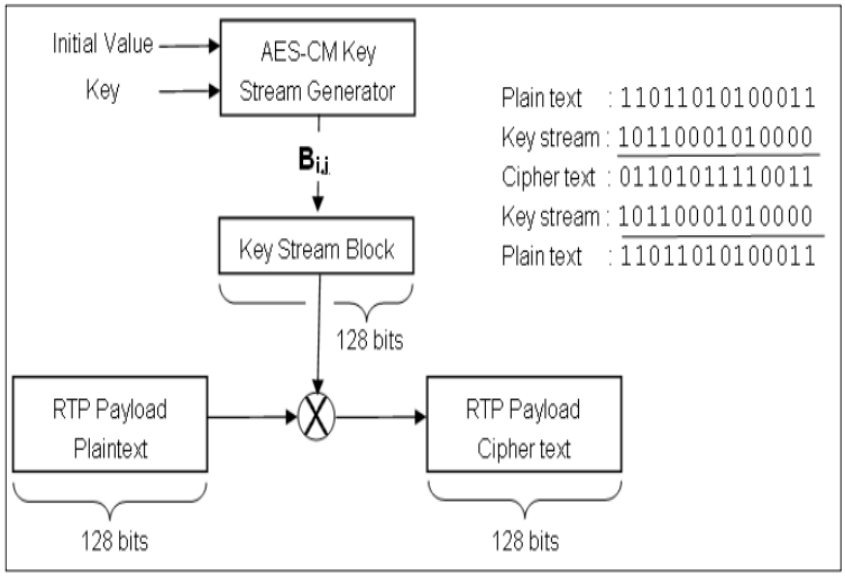


Figura 10.5: Cifratura del pacchetto SRTP

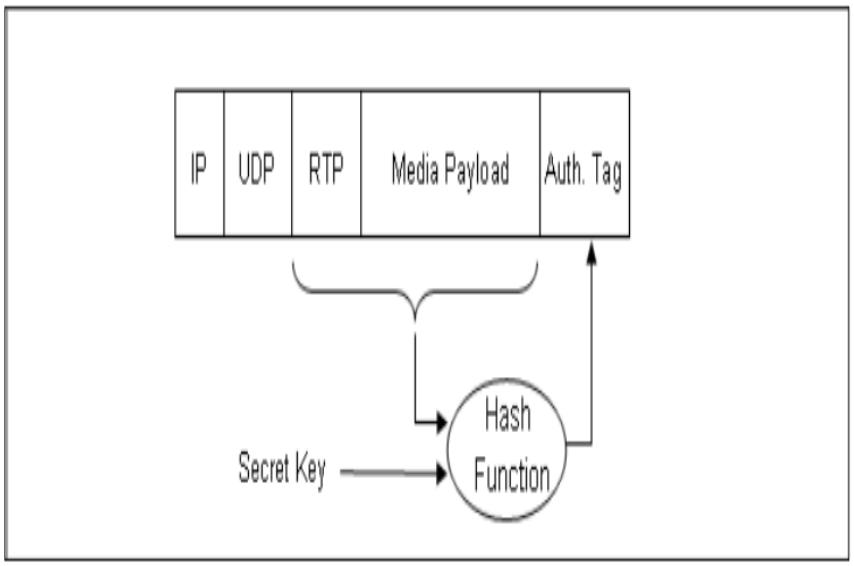


Figura 10.6: Autenticazione del pacchetto SRTP

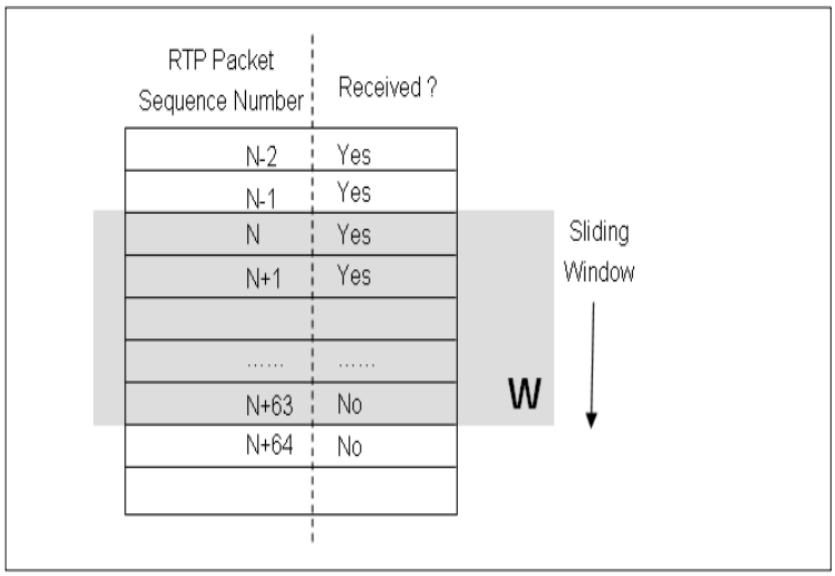


Figura 10.7: Sliding Window per la protezione da attacchi a replay

### 10.5.5 Protezione contro gli attacchi a replay

L'indice di un pacchetto SRTP può essere decifrato come valido anche se questo era invalido. Non può esserci un controllo d'integrità finchè non viene determinata la chiave di autenticazione. I tipi di attacchi utilizzati in questo caso sono counter replay attack, Rollover Counter (ROC) e sliding window. Il numero di sequenza a 16 bit dall'header RTP viene aggiunto ai 32 bit del ROC SRTP, memorizzato nel contesto crittografico, per ottenere un numero di sequenza a 48bit.

Come mostra la figura 10.7, l'indice di un pacchetto ricevuto deve essere compreso all'interno dell'intervallo della sliding window, e il bit *Received* corrispondente non deve essere settato affinchè il pacchetto possa essere passato al passo successivo. Se il pacchetto non rispetta tali criteri viene scartato. Se un attaccante sceglie un numero di sequenza casuale, e la grandezza della finestra è 64, c'è una percentuale del 99.9% $(1-64/2^{16})$ che il pacchetto sia scartato prima che siano effettuati calcoli più complessi per autenticarlo.

# Bibliografia

[1] Brian Baskin,Larry Chaffin,Michael Cross, Jan Kanclirz, Antonio Rosela, Choon Shim,Andy Zmolek (2006), *Pratical Voip Security*, Syngress, ISBN:1597490601

[2] Thomas Porter, Michael Gough (2007), *How to Cheat at Voip Security*, Syngress, ISBN:1597491691

[3] David Endler and Mark Collier (2006), *Hacking Exposed Voip: Voice Over IP Security Secrets and Solutions*, McGraw-Hill, ISBN:0072263644

[5] Diego Gosmar, Giuseppe Innamorato, Dimitri Osler, Stefano Osler (2006), *Asterisk e Dintorni. La guida italiana all'open source*, Apogeo, 9788850310418

[5] J.Arkko,V.Torvinen,G.Camarillo,Ericsson,A.Niemi,T.Haukka,Nokia (2002), *Security Mechanism Agreement for the Session Initiation Protocol*