

Architettura di conferenza SIP con spettatori multicast e supporto per streaming

Viene proposta una architettura realizzativa di un servizio di conferenza centralizzata SIP, in grado di permettere la comunicazione di gruppo tra più User Agent, e di permettere funzionalità avanzate come

- la replica della conferenza in multicast
- la distribuzione tramite live streaming
- l'immissione di contenuti preregistrati
- la registrazione della conferenza per una sua riproduzione successivamente

Sono innanzitutto definite le componenti di base necessarie alla realizzazione, come la segnalazione e la struttura del server di conferenza; quindi si passa ad illustrare l'interazione con la distribuzione multicast e con lo streaming server. Infine, si accenna a come realizzare le funzionalità *intermedie* per mezzo di un apposito componente di *middleware*.

Indice generale

Protocolli di Segnalazione.....	1
Componenti del Server di Conferenza.....	1
Integrazione dei componenti.....	1
Live Streaming.....	1
Controllo Conferenza.....	1
Streaming.....	1
Portale di Monitoraggio e Controllo.....	1
Middleware.....	1

Protocolli di Segnalazione

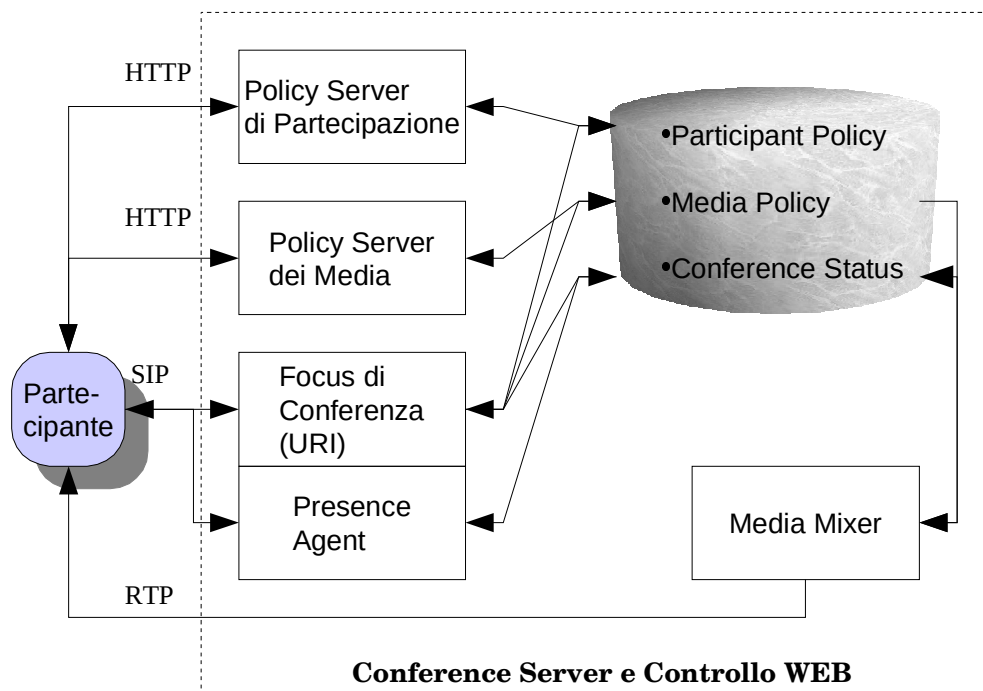
- Analisi dei requisiti - La segnalazione è responsabile di consegnare ai componenti della architettura i comandi che ne permettono l'operatività, così come di inviare/ricevere controlli con gli utenti. Si sceglie SIP come framework unitario di segnalazione, in virtù delle proprietà elencate di seguito. Nondimeno, si farà uso di HTTP per l'accesso al portale, di RTSP per la gestione dello streaming, e di SDP per esprimere le modalità di codifica e di trasporto dei media.
- SIP (Session Initiation Protocol)
 - Invito a partecipare - E' il metodo di base di SIP, mediante il quale sono trasmesse (incapsulando SDP) le informazioni necessarie a prendere parte alla conferenza, come codec audio/video, trasporto e numero di porta. Può essere inviato dal server di conferenza (SC) verso gli utenti.
 - Offerta/Risposta SDP- Mediante questa caratteristica di SDP su SIP, è possibile negoziare (tra SC ed utenti) il migliore formato di ricezione per i media.
 - Controllo di chiamata di terza parte (3pcc) - Le più recenti estensioni SIP definiscono meccanismi che permettono di deviare una connessione in corso verso soggetti differenti, sotto il controllo di una terza parte. Ciò consente di ripartire l'attuazione del servizio di conferenza tra componenti, di coordinarne il funzionamento, e di sollevare gli utenti finali dalla partecipazione attiva nei processi di call setup.

- Gestione di eventi asincroni - Un meccanismo di SUBSCRIBE/NOTIFY consente a diverse parti di mantenersi costantemente aggiornate, relativamente alle modifiche delle informazioni di stato dei diversi componenti. Ad esempio, i partecipanti alla conferenza possono essere informati automaticamente sulla composizione istantanea della conferenza, e grazie allo stesso meccanismo, è possibile rendere dinamici i contenuti pubblicati dal portale, anche in presenza di una architettura distribuita.
- HTTP - Benché non propriamente un protocollo di segnalazione, è usato per accedere alle informazioni pubblicate sul portale, nonché per impartire all'SC comandi ed azioni sempre attraverso il portale. Inoltre, può essere usato anche per il passaggio di informazioni tra le componenti del servizio, utilizzando i metodi GET/POST, in alternativa ai metodi SUBSCRIBE/NOTIFY di SIP.
- RTSP (Real Time Streaming Protocol) - E' usato per il controllo di uno Streaming Server, e può essere invocato dagli utenti per partecipare alle sessioni di conferenza in forma passiva. Allo stesso tempo, può essere usato tra componenti del servizio di conferenza, ad esempio per richiedere il playback di contenuti preregistrati all'interno di una conferenza, oppure per richiedere la registrazione di una conferenza in corso di svolgimento. In questi ultimi due casi, il controllo di playback/registrazione ha origine a seguito di una visita al portale.
- SDP (Session Description Protocol) - Permette di descrivere i media facenti parte di una stessa sessione multimediale, nei termini del tipo (audio/video/short message/chat/dati), codifica (es. GSM, AAC, AMR per audio, H261, H263, MPEG per video), indirizzi e porte (sia uni che multicast), trasporto (TCP/UDP/RTP), direzione (sendonly/rcvonly/sendrec). Grazie al suo incapsulamento nei protocolli di segnalazione e di controllo chiamata e delivery (SIP, RTSP, HTTP), costituisce il formato di rappresentazione unificante per la corretta ricezione dei media.

Componenti del Server di Conferenza

In ambito H.323, questa funzione è spesso identificata con quella di una cosiddetta MCU (Multipoint Control Unit) spesso percepita come una unità monolitica che partecipa appieno nella terminazione della segnalazione, ed assolve completamente il proprio compito con il mixing dei media audio e la combinazione del video degli ultimi quattro parlatori. Al contrario, la decomposizione della stessa in componenti, in grado di assolvere compiti ristretti, è la soluzione che permette la portabilità, il riuso, e la espandibilità verso ulteriori features, e che in sostanza permette di realizzare il sistema di conferenza che si intende descrivere qui.

In figura è illustrata una decomposizione delle funzioni del Server di Conferenza (SC) in elementi indipendenti. Una interfaccia HTTP permette di impostare le *Policy di Partecipazione* (chi ammettere se ne fa richiesta, oppure i partecipanti che debbano essere invitati) e dei *Media* (chi può parlare e chi no, quali media sono previsti, in che modo i media vengono mixati/composti). Le policy risiedono in una base dati, che è consultata dagli altri componenti del SC per interrogazione, ovvero mediante notifica SIP da parte dei componenti (esterni od appartenenti al *Middlware*) che abbiano effettuato un SUBSCRIBE allo stato della policy.



Il *Focus* costituisce la URI SIP che rappresenta la conferenza, ossia che viene invitata (o che invita) dai partecipanti, che può essere pubblicizzata con meccanismi “fuori banda”, e che determina l'ammissione alla conferenza, nei limiti imposti dalla *Policy di Partecipazione*, e con le caratteristiche definite dalla *Media Policy*; in particolare, il *Focus* si occupa di svolgere la negoziazione dei codecs audio/video mediante offerte/risposte SDP, e mantiene aggiornate le informazioni di *Conference Status* che sono usate dal *Presence Agent* per notificare i partecipanti della composizione istantanea della conferenza.

Il *Media Mixer* svolge il compito di mescolare audio e video dei partecipanti, in accordo alla *Media Policy* ed al *Conference Status*, contribuendo a mantenere aggiornato quest'ultimo, mediante l'indicazione dell'identità delle sorgenti attive.

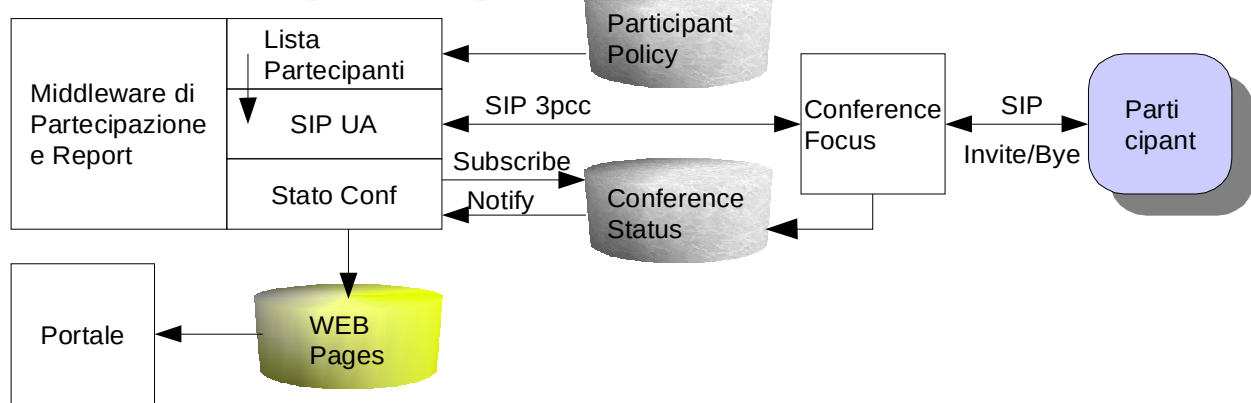
Il *Presence Agent* ha il compito di distribuire tra i partecipanti informazioni di stato, come l'elenco completo dei partecipanti, ed il nome delle sorgenti audio attive, e di quelle video che sono ritrasmesse a tutti.

Integrazione dei componenti

Allo scopo di semplificare il più possibile l'insieme delle funzionalità richieste agli user agent dei partecipanti, molto lavoro deve essere svolto “dietro il portale” da parte di un apposito MIDDLEWARE che svolge il ruolo di *Controllore di Conferenza*, e che viene diretto nelle sue operazioni mediante interfaccia WEB. In tal caso, il *Focus* e la *Presentity* dialogherebbero unicamente con il MIDDLEWARE, che si incaricherebbe di invitare i partecipanti mediante i meccanismi SIP di *Third Party Call Control*, e generare le pagine WEB relative allo stato della conferenza. Dal punto di vista sia del Server di Conferenza che degli altri partecipanti, il MIDDLEWARE di controllo è assimilabile ad un ulteriore partecipante, facente funzioni di segretario.

In presenza di User Agent evoluti, le stesse informazioni e controlli che sono mediati dal *Middleware*, possono altresì essere accedute in modo diretto, previa abilitazione espressa dalla *Participant Policy*.

Middleware di Partecipazione e Report



Live Streaming

Allo scopo di permettere la partecipazione alla conferenza del maggior numero di spettatori passivi, ovvero per i quali non è previsto che possano prendere parola, il *Media Mixer* provvede a generare uno o più ulteriori flussi RTP audio, ed ulteriori video, in base all'SDP che li descrive, e che fa parte della *Media Policy*. Tali flussi sono diretti verso uno o più indirizzi Multicast, in modo da permettere a tutti gli interessati di riceverlo.

Controllo Conferenza

Può risultare utile permettere di parlare ad una persona alla volta, ovvero ad una persona più un moderatore. Allo stesso tempo, può essere utile modificare la composizione del video trasmesso ai partecipanti, definendo composizioni ad hoc. Ciò può essere facilmente ottenuto agendo sulla *Media Policy*, da parte di un operatore umano via WEB.

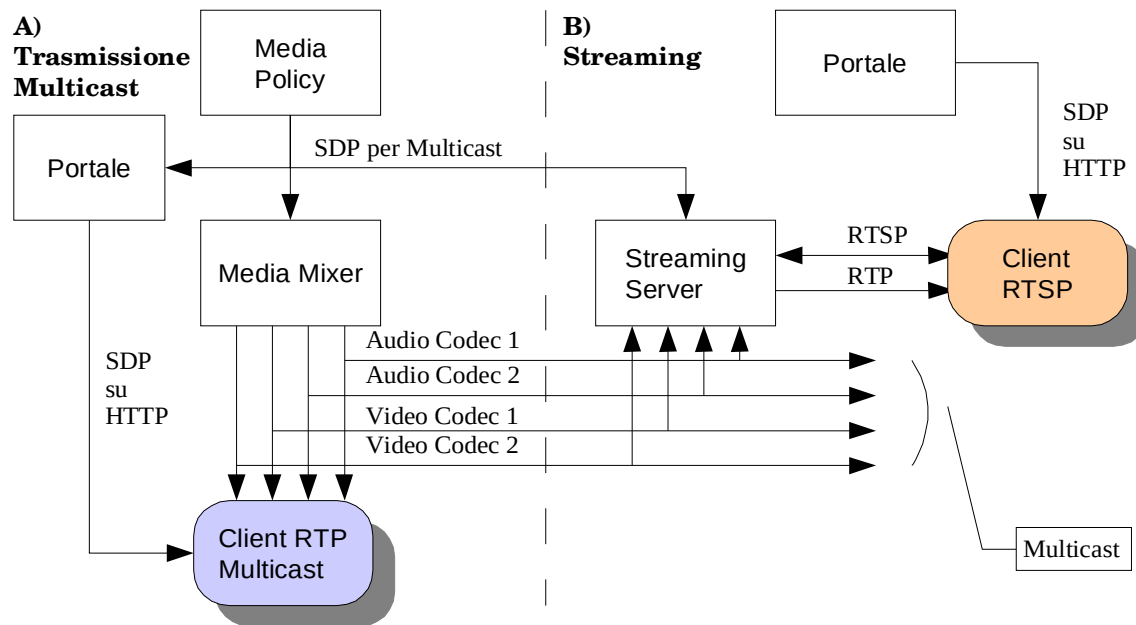
Streaming

Questa sezione affronta le seguenti problematiche:

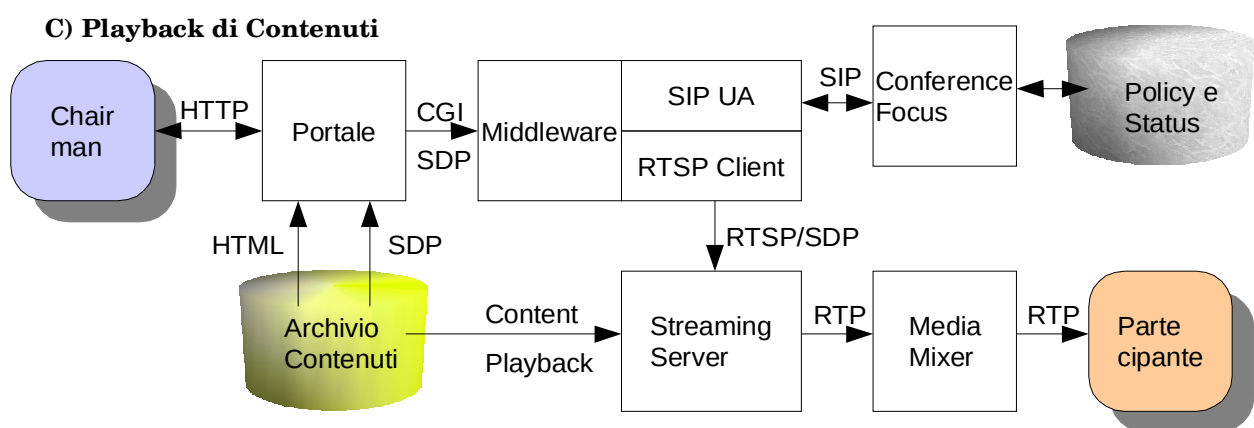
- A) - Trasmissione multicast dei media mixati
- B) - Trasmissione unicast dei media mixati
- C) - Riproduzione nella conferenza di materiale preregistrato
- D) - Registrazione della conferenza in un archivio multimediale
- E) - Directory del materiale archiviato e sua indicizzazione

Analizziamo ciascun punto. Le figure che seguono tentano di illustrare le procedure descritte.

- A) La trasmissione Multicast dell'uscita del *Mixer* audio-video è prodotta dal mixer stesso, in base alle indicazioni espresse dalla *Media Policy*. Quest'ultima contiene tra l'altro una descrizione SDP della sessione di uscita, che specifica i parametri della trasmissione (formati audio e video, anche multipli, ed indirizzi verso cui trasmettere). Lo stesso SDP è accessibile a partire dalle informazioni pubblicate sul portale, ed usato dai partecipanti passivi in grado di ricevere il multicast, mediante un opportuno client abilitato alla ricezione di stream multicast.
- B) L'erogazione Unicast dei contenuti erogati in (A) può avvenire configurando uno Streaming Server (es. Apple Darwin) in modalità RELAY, che quindi anziché servire un contenuto preregistrato, ritrasmette (replicandolo per ogni client) lo stesso flusso RTP che è erogato in Multicast; qualora questo sia disponibile in più formati, occorre prevedere un diverso SDP per ogni formato (o sua combinazione). L'uso di RTSP per il controllo dello streaming da parte dei client, permette di sospendere temporaneamente la ricezione dei contenuti, in modo indipendente per ogni partecipante.



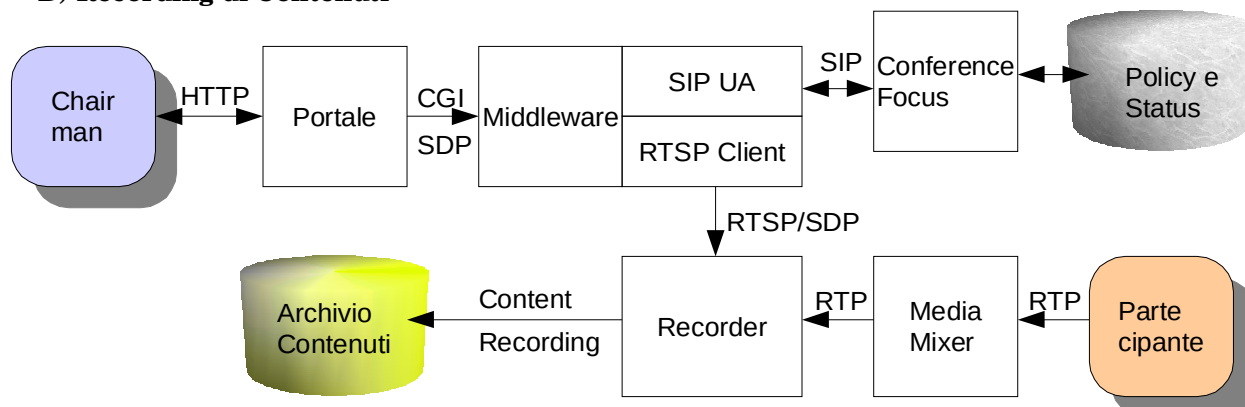
C) Durante una conferenza, può rendersi necessario sottoporre all'attenzione dei partecipanti un contenuto multimediale preesistente, e memorizzato in un archivio al quale lo streaming server ha accesso diretto. In tal caso, deve essere possibile realizzare una serie di azioni formalmente equivalenti ad *invitare* (via SIP) uno Streaming Server a trasmettere il contenuto verso la conferenza, alla stessa stregua di un partecipante qualunque. La realizzazione di questo servizio in modo semplice, prevede la possibilità di selezionare il contenuto, e di iniziare/sospendere/terminare il playback, agendo su di una apposita pagina del portale, da parte di un *Chairman*. Tali azioni determinano l'invocazione di primitive di servizio offerte dal *Middleware*, che a sua volta genera comandi SIP verso il *Conference Focus* (per aggiungere lo SS alla conferenza) e comandi RTSP verso lo SS. Uno stesso SS può iniettare contenuti in diverse conferenze.



D) La registrazione di una conferenza può costituire una preziosa risorsa nel caso in cui la conferenza stessa rivesta aspetti didattici o di formazione, che possono così essere fruiti successivamente, ed eventualmente consultati a richiesta. L'architettura necessaria a realizzare questa funzione ricalca quella esposta al punto C), prevedendo ancora un *Middleware* invocato dal Chairman per mezzo del portale, che provvede ad inserire (via SIP) nella conferenza una entità RECORDER, ed a comandare alla stessa (via RTSP) la registrazione nell'archivio dei

media specificati mediante SDP.

D) Recording di Contenuti



E) Per la corretta realizzazione dei punti C) e D), occorre definire le modalità di archiviazione ed indicizzazione dei contenuti, tali da permetterne un recupero agevole, ed una facile ricerca. Allo scopo si prevede la formalizzazione di *metadata* XML, capitalizzando gli sviluppi proposti dai gruppi di lavoro che operano in ambito internazionale (es. TF-NETCAST di TERENA).

Portale di Monitoraggio e Controllo

La realizzazione delle funzioni di controllo conferenza e streaming mediante una interfaccia unificata in un portale, consente di limitare al massimo il livello di competenza tecnica richiesto ai partecipanti, che possono usare applicazioni/client di tipo diverso, e non particolarmente mirate al conferencing, come Endpoints H323 o User Agent SIP. Le pagine del portale dovranno consentire di

- Creare un *Identificativo di Conferenza* (SIP URI), descriverne il tema, delimitarne la durata
 - ◆ Definire la *Policy di Partecipazione* (broadcast, meeting, gruppo chiuso, multicast, streaming, media recording)
 - ◆ Definire la *Media Policy* (codecs supportati, modalità di composizione video, indirizzi, erogazione multicast e streaming)
- Visualizzare le conferenze in corso, e di prossima attivazione
- Visualizzare le informazioni relative allo status di ogni conferenza in corso:
 - ◆ i partecipanti presenti, con che media/codec, attivi/passivi
 - ◆ le policy in vigore per la conferenza
 - ◆ gli indirizzi multicast e/o RTSP per assistere passivamente
 - ◆ Fornire modalità di gestione della conferenza:
 - ★ accordare/revocare la parola agendo dinamicamente sulle policy
 - ★ eseguire inviti di massa di un gruppo chiuso di utenti
 - ★ escludere esplicitamente partecipanti selezionati
 - ★ invocare il playback di contenuti preesistenti
 - ★ controllare la registrazione di una conferenza in corso
 - ★ terminare la conferenza in modo prematuro

- Accedere ai contenuti preregistrati mediante chiavi di ricerca e directory

Middleware

Per realizzare le funzionalità descritte, occorre sviluppare un componente di *Middleware* che integri l'accesso alle basi dati, con l'esecuzione di agenti SIP che interagiscano con gli altri componenti. In particolare, possiamo individuare i seguenti componenti necessari a realizzare le funzioni di

- Controllo partecipazione e Report:
 - ◆ accesso alla *Participant Policy*;
 - ◆ esecuzione di uno UA SIP con capacità di eseguire comandi 3pcc con il *Focus*;
 - ◆ esecuzione di uno UA SIP con capacità di sottoscrizione alle informazioni di *Conference Status*;
 - ◆ generazione dinamica di pagine HTML contenenti lo stato istantaneo dei componenti la conferenza, da visualizzare sul *Portale*.
- Controllo Playback:
 - ◆ esecuzione di uno UA SIP per l'ammissione del *Playback* in conferenza;
 - ◆ esecuzione di un client RSP per il controllo dello SS che effettua il *Playback*.
- Controllo Recording:
 - ◆ esecuzione di uno UA SIP per l'ammissione del *Recorder* in conferenza
 - ◆ esecuzione di un client RSP per il controllo del *Recorder*