



**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET  
POPULAIRE**

**MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE  
RECHERCHESCIENTIFIQUE**

**UNIVERSITE DE KASDI MERBAH OUARGLA**

Faculté des Nouvelles Technologies de l'Information et la  
Communication

Département de l'informatique et Technologies de l'information

N° d'ordre :

N° de série :

**Mémoire**

**MASTER ACADEMIQUE**

**Domaine : Math et informatique**

**Filière : Informatique**

**Présenté par**

**- SAADI Sami**

**- KEMOUKH Djawhara**

**Spécialité: Master Informatique industriel**

**Thème**

***Réalisation d'une application web de visio-conférence multipoints***

Soutenu publiquement

**Devant le jury :**

**Président : Mr. BENKADOUR Mohamed Kamel**

**Encadreur : Mr. KAFI Redouan**

**Co-Encadreur : Mr. MEZATI Messaoud**

**Examineur : Mr. EUSHI Saleh**

**Année Universitaire : 2015/2016**

# Table des matières

Liste des Figures.....	4
Résumé .....	5
Introduction .....	6
Chapitre I : Introduction à la visio-conférence .....	8
Introduction.....	9
1.    Qu'est-ce que la vidéoconférence ? .....	9
2.    Principe de fonctionnement .....	9
3.    Vidéoconférence point-à-point.....	10
4.    Vidéoconférence multipoint .....	12
5.    La langue de la vidéoconférence .....	14
6.    Norme H.323 .....	14
a.    Norme H.245 .....	15
b.    Norme H.225 .....	15
c.    Norme H.264 .....	16
d.    Advanced Audio Coding .....	18
7.    SIP (Session Initiation Protocol) .....	19
8.    Architecture et fonctionnalités des ponts .....	21
9.    Gestion des appels vidéo et conversion de protocoles .....	23
10.   Gestion des périphériques.....	24
11.   Sécurité .....	24
12.   Gestion du contenu .....	25
Conclusion .....	26
Chapitre II : Conception de l'application.....	27
Introduction.....	28
1.    Diagramme de classe .....	28
2.    Diagramme d'activité .....	29
3.    Choix des outils .....	29
a.    Codec Vidéo .....	29
b.    Codec Audio.....	30
Conclusion .....	31

Chapitre III : Implémentation de l'application .....	32
Introduction.....	33
1. Langage de programmation.....	33
2. Transmission.....	35
3. Architecture de l'application .....	35
4. Performance.....	39
Conclusion .....	41
Conclusion.....	42
Bibliographie.....	<b>Erreur ! Signet non défini.</b>

# Liste des Figures

Figure I.1 Solutions portables de visio-conférence .....	10
Figure I.2 Solution intégré dans la salle .....	10
Figure I.3 Frames d'une vidéo en format H.264.....	16
Figure II.1 diagramme de classe .....	28
Figure II.2 Diagramme d'activité .....	29
Figure III.1 Comparaison de performance entre PHP et ASP.NET MVC .	34
Figure III.2 Architecture de l'application .....	35
Figure III.3 Demande d'autorisation.....	36
Figure III.4 La capture de la caméra est affichée .....	37
Figure III.5 Liste des salles publiques .....	37
Figure III.6 L'utilisateur entre son nom pour rejoindre la salle .....	38
Figure III.7 L'échange des flux Vidéo et Audio .....	38
Figure III.8 Utilisation de la bande passante de l'application .....	39
Figure III.9 Utilisation de la bande passante de Skype.....	40
Figure III.10 Utilisation de la bande passante de Facebook .....	40

# Résumé

Les entreprises réparties sur des sites distants utilisent la visioconférence pour réduire les coûts de déplacement, tout en ayant des réunions fréquentes.

Dans ce travail nous allons réaliser la mise en œuvre d'une application de visioconférence multipoints basé sur le Web, cette application sera implémentée en utilisant la plateforme ASP.NET MVC 6. Nous avons choisis le standard H.323 avec un encodage vidéo H.264 et un encodage audio AAC-LC, le travail visera à donner une performance optimale de consommation de bande passante et de préservation de la qualité d'image et de son.

**Mots clés :** Visio-conférence, Web, H.323

## *Abstract*

Companies spread over remote locations using videoconferencing to reduce travel costs while having frequent meetings.

In this work we will carry out the implementation of an application for web-based multipoint videoconferencing, this application will be implemented using the ASP.NET MVC 6 platform. We chose the H.323 standard with H.264 video Codec and AAC-LC audio Codec. The work will aim to give optimum performance of bandwidth consumption while preserving the quality of picture and sound.

**Key words:** Videoconferencing, Web, H.323

# **Introduction**

La Visio téléphonie est longtemps restée un serpent de mer, déjà montrée aux expositions universelles de Bruxelles en 1958 et de Montréal en 1967, ainsi que mise en place de façon expérimentale (et coûteuse) à Biarritz en 1984. Cette stagnation venait du faible débit des lignes de communication jusqu'à l'utilisateur final. Plusieurs facteurs d'évolution ont permis de mettre en œuvre ce service : les réseaux numériques de bout en bout, les baisses de coût dans les équipements comme les caméras vidéo, mais surtout les progrès dans les techniques de codage (compression des données) audio et vidéo. Les premières liaisons universitaires, en France, datent de 1987.

On nomme visioconférence la combinaison de deux techniques :

- La visiophonie ou vidéotéléphone, permettant de voir et dialoguer avec son interlocuteur ;
- La conférence multipoints ou conférence à plusieurs, permettant d'effectuer une réunion avec plus de deux terminaux.

Dans la pratique, le terme reste toutefois utilisé même lorsque les interlocuteurs ne sont que deux. [7]

L'objectif de ce travail est la réalisation d'une application de visioconférence basé sur le web en utilisant la norme H.323 de visioconférence sur IP.

Nous avons réparti ce travail en 3 chapitres :

Dans le premier chapitre nous allons donner une introduction à la visioconférence, les outils logiciels et matériel nécessaire à l'établissement d'une réunion en vidéo, les standards les plus utilisés dans cette industrie et l'architecture des systèmes de visioconférence.

## *Introduction*

---

Ensuite dans le chapitre suivant nous présentons notre implémentation du système de visio-conférence sous forme d'une application web ASP.NET MVC, qui utilise WebRTC et SignalR pour établir la communication temps réel.

# **Chapitre I : Introduction à la** **visio-conférence**



## **Introduction**

Au cours des prochaines années, nous verrons une croissance explosive dans l'utilisation de la vidéoconférence comme un outil fondamental pour les entreprises pour améliorer la communication et la collaboration entre les employés, les partenaires et les clients. La technologie a considérablement évolué depuis les premiers adoptants à sa forme actuelle déployement en masse sur le marché. Les statistiques indiquent que la moitié des travailleurs de l'information ont un certain type de solution de vidéo personnelle en 2016. Avec la vidéoconférence devenant une composante essentielle de l'infrastructure informatique qui permet la communication et la collaboration, les entreprises seront à la recherche des fournisseurs de téléphonie, des applications d'entreprise et des services d'infrastructure de réseau afin d'inclure cette fonctionnalité dans le cadre de leur offre.

Ce chapitre examine les composantes de base de la technologie, et les considérations pour le déploiement de solutions de vidéoconférence.[10]

### **1. Qu'est-ce que la vidéoconférence ?**

Pour définir les bases pour l'élaboration future, au niveau le plus simple, une conférence vidéo est une réunion en ligne (ou une réunion sur la distance) qui a lieu entre deux parties, où chaque participant peut voir une image de l'autre, et où les deux parties sont capable de parler et d'écouter les autres participants en temps réel. [10]

### **2. Principe de fonctionnement**

Les composants nécessaires pour établir une vidéoconférence comprennent :

- Un microphone, webcam et haut-parleurs.
- Un écran.
- Un programme qui capture le flux vocal à partir du microphone, encode, transmet à l'autre participant, et en même temps décode le flux vocal

numérique étant reçu du participant distant à la conférence vidéo (plus communément appelé un «Codec»).

- Un logiciel qui relie les deux parties ensemble à travers une connexion numérique, la gestion de l'échange de voix et vidéo entre les participants. À chaque extrémité de la connexion, le trafic vidéo et de la voix est combiné et remis à chaque participant sous la forme d'une image vidéo et un flux audio en temps réel.
- Un outil de gestion optionnel pour la programmation des sessions de visioconférence.

À un niveau légèrement plus avancée, il est également possible de fournir la possibilité de partager le contenu d'un dispositif au cours d'un appel vidéo. La qualité et le type de contenu qui peut être partagé dépend du taux d'échange de données pendant l'appel.

La terminologie utilisée par les utilisateurs de vidéoconférence pour décrire le processus de numérotation dans et en participant à une réunion virtuelle est connu comme « se joindre à un pont ». Différentes salles de réunion virtuelle sont affectées des « numéros de ponts», et les utilisateurs se joint à un appel vidéo en « composant un numéro de pont ».[10]

### **3. Vidéoconférence point-à-point**

Les réunions vidéo se produisent de deux manières distinctes : soit en point à point ou multipoint. En point-à-point, le scénario le plus simple est l'endroit où une personne ou un groupe est connecté à un autre. Les composants physiques (microphone et caméra) qui permettent la réunion d'avoir lieu sont souvent intégrés dans des solutions informatiques de bureau comme un ordinateur portable ou une tablette, ou peuvent être incluse dans des solutions hardware intégré dans la salle.

Les solutions de bureau ont tendance à être utilisées par des particuliers, les solutions intégrées dans la salle utilisent la technologie de visioconférence dédiée où des groupes de personnes peuvent être vus, entendus et peuvent naturellement participer à la réunion.



**Figure I.1** Solutions portables de visio-conférence [3]



**Figure I.2** Solution intégré dans la salle [3]

#### **4. Vidéoconférence multipoint**

Dans les appels vidéo multipoints, trois ou plusieurs emplacements sont reliés entre eux, où tous les participants peuvent voir et entendre les uns des autres, ainsi que de voir tout contenu partagé au cours de la réunion.

Dans ce scénario, les flux d'information numérique de la voix et de la vidéo et le contenu sont traités par un logiciel central, indépendant. Après la combinaison de la vidéo et de la voix du trafic du participant, le programme ré-envoie un flux de données collective retour participants à la réunion sous la forme d'audio en temps réel et des images vidéo.

Les individuels peuvent participer à une réunion dans un mode "audio seulement", ou combiner l'audio avec des images vidéo de la réunion sur l'écran. En fonction de la capacité technique du système de vidéoconférence utilisé, les images vues par les participants sont soit classés comme «interlocuteur actif» ou «présence continue».

En mode «interlocuteur actif », l'écran ne fournit qu'une image de la personne qui parle à tout moment. Dans des solutions plus avancées avec le mode «présence continue», le pont divise l'image sur l'écran en un certain nombre de domaines différents. La personne qui parle à tout point dans le temps est présenté dans une grande zone centrale, les autres participants à la réunion sont affichés autour de l'image centrale. Le mode «présence continue» permet ainsi aux participants de la réunion de visualiser et d'interagir avec tous les participants dans une «salle de réunion virtuelle.

Le logiciel qui crée la «salle de réunion virtuelle» et le matériel de traitement numérique sur lequel il réside, sont souvent appelé un pont vidéo, ou «pont», pour faire court. Un autre terme pour un pont qui est souvent utilisé est une « unité de contrôle multipoint » vidéoconférence ou « MCU ». [10]

Alors que la vidéoconférence point à point est relativement simple, la création et la gestion des conférences vidéo multipoints peuvent être complexes. Un MCU doit être en mesure de créer, contrôler et faciliter plusieurs réunions simultanées de vidéoconférence en direct. Une complexité supplémentaire est ajoutée lorsque différents endroits peuvent se connecter à la réunion sur les flux numériques ou analogiques à des vitesses différentes, avec le transport de données différentes et des protocoles de signalisation utilisés pour faciliter la communication.

Pour relier ces utilisateurs en une réunion virtuelle commune, le MCU doit donc être en mesure de comprendre et de traduire entre plusieurs protocoles différents (ex H.264 pour la communication sur IP et H.263 pour le RNIS). Le MCU permettra également ceux qui rejoignent le pont vidéo de le faire à la vitesse la plus élevée et la meilleure qualité possible que leur système individuel peut prendre en charge. Bien qu'il existe deux processus distincts qui se déroulent ici, ceci est souvent appelé conjointement "transcodage".

Il est important de noter que pas tous les ponts offrent une telle capacité de transcodage, et cela peut sérieusement affecter la qualité et l'expérience des appels vidéo. Lorsque le transcodage n'est pas fournies et les utilisateurs composant un numéro de pont avec des vitesses de connexion différentes, il est possible que le pont ne peut être en mesure de soutenir la réunion vidéo en établissant les connexions au niveau du plus petit dénominateur commun. Pour illustrer l'effet négatif, envisager une réunion qui a lieu avec la plupart des utilisateurs se connectant au pont à partir du réseau d'entreprise à grande vitesse, mais où une ou deux personnes cadran à la réunion de la maison sur DSL à faible bande passante ou RNIS. Dans ce cas, l'expérience des nombreux utilisateurs de l'entreprise est rétrogradée au plus petit dénominateur commun des utilisateurs à domicile, ce qui pourrait rendre l'appel vidéo inefficace. Lorsque le transcodage efficace est pris en charge par le MCU, ceux sur le

réseau de l'entreprise va continuer à profiter de la qualité vidéo HD, tandis que les utilisateurs distants reçoivent la qualité en rapport avec leurs vitesses de connexion. [10]

En résumé, quand un MCU est bien conçu, intégrant facilement avec plusieurs fournisseurs et permettant aux utilisateurs d'appeler au débit de données et de la résolution qu'ils veulent ou ont besoin, le résultat est une expérience transparente facile pour tous les utilisateurs, permettant aux gens de se concentrer sur la réunion, pas la technologie.

## **5. La langue de la vidéoconférence**

Comme la technologie de vidéoconférence a évolué, deux protocoles principaux ont émergé pour fournir le contrôle de signalisation pour l'établissement, le contrôle et la terminaison des appels de vidéoconférence : SIP (Session Initiation Protocol) et H.323.

Pour le codage et le décodage de l'information visuelle, l'industrie se dirige vers la norme de l'industrie connue sous le nom H.264, qui a été développé pour fournir une vidéo de haute qualité à faible bande passante sur une large gamme de réseaux et de systèmes. Une extension du protocole H.264 est Scalable Video Coding (SVC), qui est mis en place pour faciliter l'activation de la vidéoconférence sur un large éventail de dispositifs, tels que les tablettes et les téléphones mobiles.[10]

## **6. Norme H.323**

La norme H.323 fournit une base pour les communications audio, vidéo et de données à travers les réseaux protocole Internet (IP). H.323 est une recommandation de l'Union Internationale des Télécommunications (UIT) qui établit des normes pour les communications multimédias sur des réseaux locaux (LAN). La norme H.323 fait partie d'un plus large éventail de normes de vidéoconférence (H.32x) pour la vidéoconférence sur différents supports de

réseau. Par exemple, H.320 prend en charge la vidéoconférence via le réseau numérique à intégration de services (RNIS), H.321 prend en charge la vidéoconférence sur le mode de transfert asynchrone (ATM), H.324 prend en charge la vidéoconférence sur la norme Plain Old Telephony Services (POTS), et H.323 prend en charge la vidéoconférence sur les réseaux locaux IP.[2]

La spécification H.323 se compose de plusieurs protocoles, y compris :

**a. Norme H.245**

Fournit la signalisation de commande utilisée pour échanger des messages de contrôle de bout en bout. Ces messages de commande comportent des informations concernant :

- L'échange de capacités
- Ouverture et fermeture des voies logiques utilisées pour transporter des flux média.
- messages de contrôle de flux.
- Les commandes générales et indications. [2]

**b. Norme H.225**

Fournit l'enregistrement, l'admission et le statut (RAS), qui est le protocole utilisé entre les dispositifs H.323 et le portier pour l'enregistrement de l'appareil. Le protocole RAS est utilisé pour effectuer l'enregistrement, le contrôle d'admission, les mises à jour d'utilisation de bande passante, le statut et les procédures de désengagement entre les dispositifs H.323 et le portier. H.225 est également utilisé lors de la configuration d'appel pour ouvrir un canal de signalisation d'appel en utilisant le protocole standard de messagerie Q.931. [2]

Standard	Fonctions supportées
H.225	RAS, démarrage et arrêt forcé des appels, établissement des appels Q.931
H.245	Messagerie de control d'appels

H.261	Format vidéo
H.263	
H.264	
G.711	Format audio
G.722	
G.723	
G.728	
AAC-LC	
AAC-LD	

**Table I.1** Protocoles pris en charge par la norme H.323

**c. Norme H.264**

Ce standard a été développé en réponse à la nécessité croissante d'une compression plus élevée des images animées pour diverses applications telles que la vidéoconférence, les supports de stockage numérique, la radiodiffusion télévisuelle, internet streaming, et de la communication. Il est également conçu pour permettre l'utilisation de la représentation vidéo codé de manière flexible pour une grande variété d'environnements de réseau. L'utilisation de H.264 permet à la vidéo d'être manipulé comme une forme de données informatiques et à être stocké sur différents supports de stockage, transmises et reçues sur les réseaux existants et futurs et distribué sur les canaux de radiodiffusion existants et futurs.

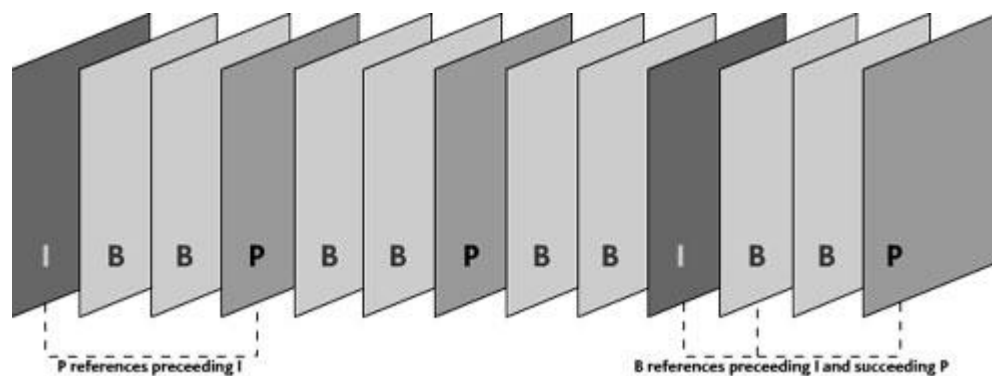
H.264 / AVC est le codec vidéo le plus récent largement utilisé par les systèmes de visioconférence actuels.

En 2001, le Motion Picture Experts Group (MPEG) d'ISO a reconnu le potentiel du développement H.263 de l'UIT-T et a formé le Joint Video Team (JVT) qui comprenait des experts de MPEG et VCEG. Le résultat est deux normes identiques : ISO MPEG4 Part 10 et de l'UIT-T H.264, avec le nom officiel Advanced Video Coding (AVC).[5]



Il y a peu de différence fonctionnelle entre les éléments de H.264 et ceux des normes H.261 et H.263 antérieures. Les changements qui font faire mentir différence principalement dans le détail au sein de chaque élément, la façon dont l'algorithme est mise en œuvre et si elle est effectuée dans le matériel ou le logiciel.

La technique de base de la prédiction de mouvement fonctionne en envoyant un frame complet suivi d'une séquence de frames qui ne contiennent que les parties de l'image qui ont changé. Les frames complets sont également connus comme des «images clés» ou «I-frames» et les frames prédits sont appelés «P-frames». Comme un frame perdue ou diminué peut causer une séquence de trames envoyées après d'être illisible, de nouvelles «I frames» sont envoyés après un nombre prédéterminé de« P-frames ». Il est la combinaison de la compression avec perte et la prédiction de mouvement qui permet aux systèmes H.261, H.263 et H.264 d'atteindre la réduction nécessaire des données tout en fournissant une qualité d'image acceptable. [1]



**Figure I.3** Frames d'une vidéo en format H.264.[13]

Avec des centaines d'experts impliqués dans la création H.264, il y avait beaucoup d'options. Certains étant plus simple et immédiatement mis en œuvre, tandis que d'autres étaient beaucoup plus complexes, mais toujours inclus. Ainsi H.264 a été organisé en quatre profils; Baseline, Extended, Main et haut. La ligne de base est la plus simple et utilise 4: 2: 0 chrominance échantillonnage et

divise l'image en blocs de 4x4 pixels, le traitement de chaque bloc séparément. Baseline utilise Universal Longueur Codage Variable (UVLC) et Context Adaptive Variable Length Coding (CAVLC) techniques qui ont un grand impact sur la bande passante du réseau. Pratiquement tous les fournisseurs prennent en charge H.264 Baseline et certains soutiennent désormais également H.264 High Profile.

H.264 est destinée à couvrir une large gamme d'applications De contenu vidéo, y compris mais sans s'y limiter à ce qui suit :

- CATV TV par câble sur les réseaux optiques, le cuivre, etc.
- Services vidéo par satellite de radiodiffusion directe DBS
- Services vidéo de ligne d'abonné numérique DSL
- DTTB Radiodiffusion numérique de télévision terrestre
- Médias ISM Interactive de stockage (disques optiques, etc.)
- MMM Multimédia postale
- Services MSPN multimédia sur des réseaux de paquets
- RTC services conversationnels en temps réel (visioconférence, visiophone, etc.)
- RVS surveillance vidéo à distance
- SSM supports de stockage Serial (VTR numérique, etc.)[4]

#### **d. Advanced Audio Coding**

(AAC, « Encodage Audio Avancé » en français) est un algorithme de compression audio avec perte de données ayant pour but d'offrir un meilleur rapport qualité sur débit binaire que le format plus ancien MPEG-1/2 Audio Layer 3 (plus connu sous le nom de MP3). Pour cette raison, il a été choisi par différentes firmes comme Apple ou Real Networks.

Les améliorations apportées par l'AAC par rapport au MP3 touchent notamment :

- les fréquences d'échantillonnage : de 8 Hz à 96 kHz (MP3 officiel : 16 à 48 kHz) ;
- le nombre de canaux : jusqu'à 48 ;
- une efficacité de codage améliorée pour les signaux stationnaires (la taille des blocs a été portée de 576 à 1 024 échantillons) ;
- une efficacité de codage améliorée pour les signaux de transition (la taille des blocs a été réduite de 192 à 128 échantillons) ;
- une meilleure gestion des fréquences au-delà de 16 kHz ;
- une jonction stéréo (*joint-stereo*) plus souple (distincte pour chaque bande) ;
- une compatibilité de gestion des droits numériques (*DRM*) pour contrôler l'utilisation des fichiers dans ce format ;
- l'ajout de nouveaux modules de compression qui peuvent être combinés pour constituer des profils.

## **7. SIP (Session Initiation Protocol)**

SIP est un protocole de commande de couche application qui peut établir, modifier et terminer des sessions multimédia (conférences) telles que des communications téléphoniques par l'Internet. SIP peut aussi inviter des participants à des sessions déjà existantes, telles que des conférences en multidiffusion. Des supports peuvent être ajoutés (et retirés) à une session existante. SIP prend en charge de façon transparente la transposition de nom et les services de redirection, ce qui sert de support à la mobilité personnelle. Les utilisateurs peuvent conserver une identification unique vue de l'extérieur, indépendamment de leur localisation dans le réseau.

SIP prend en charge cinq facettes de l'établissement et de la terminaison de communications multimédia :

- Localisation de l'utilisateur : détermination du système terminal à utiliser pour la communication.
- Disponibilité de l'utilisateur : détermination de la volonté de l'appelé à s'engager dans une communication.
- Capacités de l'utilisateur : détermination du support et des paramètres de support à utiliser.
- Etablissement de session : "sonnerie", établissement des paramètres de session à la fois chez l'appelant et l'appelé.
- Gestion de session : y compris le transfert et la terminaison des sessions, la modification des paramètres de session, et l'invocation des services.

SIP ne fournit pas de services. Plutôt, SIP fournit des primitives qui peuvent être utilisées pour mettre en œuvre différents services. Par exemple, SIP peut localiser un utilisateur et livrer un objet opaque à l'endroit où il se trouve. Si cette primitive est utilisée pour délivrer une description de session écrite, par exemple, en SDP, les points de terminaison peuvent se mettre d'accord sur les paramètres d'une session. Si la même primitive est utilisée pour livrer une photo de l'appelant aussi bien que la description de session, un service d'"ID d'appelant" peut facilement être mis en œuvre. Comme le montre cet exemple, une seule primitive est normalement utilisée pour fournir plusieurs services différents.

SIP n'offre pas de services de contrôle de conférence du genre de la commande de salle ou des votes et n'a aucune exigence sur la façon dont une conférence doit être gérée. SIP peut être utilisé pour initialiser une session qui utilise un autre protocole de contrôle de conférence. Comme les messages SIP et les sessions qu'ils établissent peuvent passer à travers des réseaux entièrement

différents, SIP ne peut pas fournir, et ne fournit pas, de capacités de réservation de ressources de réseau d'aucune sorte. [11]

## **8. Architecture et fonctionnalités des ponts**

Comme nous l'avons vu plus haut, l'association du logiciel et du matériel qui composent les salles de réunion virtuelles s'appelle « pont vidéo ». Les salles de réunion virtuelles sont désignées par leur « numéro de pont ». Lorsque plusieurs appels se déroulent simultanément, le logiciel analyse tous les flux de données à leur arrivée dans les processeurs du pont et les répartit en fonction de cette analyse.

Au niveau le plus simple, la charge de traitement des ponts dépend de quatre facteurs :

- Le nombre de sites connectés à chaque pont ;
- Le nombre de conférences téléphoniques gérées simultanément par chaque pont ;
- La quantité de données reçues dans chaque flux numérique : des résolutions audio et vidéo élevées, telles que la HD, génèrent plus de données à traiter ;
- Le degré de transcodage requis de la part du pont, qui doit gérer des appels dont les débits et les protocoles sont différents.

Plus la charge de travail s'intensifie, plus la quantité de données à traiter par chaque pont s'accroît. Dans ce cas, pour améliorer les performances, il suffit d'augmenter le nombre de processeurs de signalisation numérique utilisés pour décoder et encoder les flux numériques qui entrent et sortent des MCU. Si le pont ne parvient plus à gérer la charge de travail, les informations audio et vidéo peuvent se perdre, introduisant une latence dans les appels, ce qui risque de nuire à l'expérience de la réunion vidéo.

Pour doter le pont de ressources de traitement supplémentaires, vous pouvez soit utiliser un pont plus puissant (avec plus de processeurs), soit adopter une approche de logiciel virtuel selon laquelle le logiciel qui contrôle la fonction de signalisation fonctionne indépendamment du matériel physique.

Il n'est pas nécessaire qu'une conférence téléphonique à laquelle un numéro a été attribué ait lieu ou soit traitée par un composant matériel dédié. L'appel peut être « virtualisé » et affecté à n'importe quel pont physique disposant des ressources ou de la capacité adéquates pour traiter l'appel. Un gestionnaire de virtualisation détecte le pont physique qui offre une telle capacité et y affecte les appels entrants.

Dans les cas extrêmes (qui sont rares), le gestionnaire de virtualisation peut répartir les ressources entre plusieurs ponts physiques qui fonctionnent en tandem. Le logiciel indique alors aux ressources du pont physique de fonctionner sur un mode « parent-enfant », dans lequel un pont est « propriétaire » de la conférence téléphonique, tandis que les autres se partagent la charge de travail. Cette configuration est appelée « mise en cascade automatique ».

En mode présence continue, le pont fournit automatiquement les modèles d'écran qui seront utilisés pour afficher les participants à la réunion. Le pont peut également offrir des fonctionnalités administratives telles que l'attribution de mots de passe pour participer aux réunions, ainsi qu'une fonction IVR (Interactive Voice Response pour réponse vocale interactive) qui permet d'accueillir les participants et de leur fournir des instructions à l'aide de messages audio personnalisés.

Même si la plupart des participants se connectent eux-mêmes aux réunions de visioconférence, le pont peut être programmé pour contacter automatiquement les sites participants et les connecter, toujours automatiquement, à la réunion. Par exemple, le pont peut « réveiller » automatiquement les caméras situées

dans des salles de réunion distantes et relier ces salles à un appel planifié. Dans ce cas, les participants à ce type de réunion n'ont plus qu'à prendre place dans la salle à l'heure prévue.

## **9. Gestion des appels vidéo et conversion de protocoles**

Pour former une architecture évolutive, la plate-forme logicielle doit fournir une fonctionnalité de signalisation des appels et gérer de manière dynamique la configuration et la maintenance d'un grand nombre d'appels vidéo. L'architecture logicielle doit être capable de se reconfigurer elle-même ainsi que ses ressources en temps réel, afin que ces dernières soient utilisées au maximum de leurs capacités. Elle doit aussi être capable d'identifier les besoins en bande passante de chaque appel, la stratégie associée (priorité et importance des appels), ainsi que l'emplacement géographique des participants. De cette manière, la plate-forme logicielle peut utiliser les ressources locales au lieu de rediriger les flux de données et la signalisation d'appel vers des ressources distantes, une approche qui monopoliserait une part importante de la bande passante sur des liaisons WAN très coûteuses.

La plate-forme logicielle doit également être capable de détecter instantanément toute défaillance au niveau des ressources matérielles ou toute perte de communication sur les liaisons de l'infrastructure. Le cas échéant, le trafic est alors redirigé et les appels rétablis à l'aide d'autres ressources, sans que cela n'affecte outre mesure leur qualité.

Lorsque plusieurs systèmes, installés sur des sites client différents, tentent de se connecter au même appel vidéo à l'aide de périphériques qui utilisent des protocoles différents (par exemple H.323, RTV ou SIP), la plate-forme de visioconférence doit dans un premier temps convertir les protocoles dans un langage commun afin que l'infrastructure puisse comprendre les informations et les traiter comme il convient. En d'autres termes, la plateforme logicielle doit

établir une passerelle intrinsèque entre les périphériques qui utilisent des langages différents.

## **10. Gestion des périphériques**

Dans les déploiements et la gestion de solutions de visioconférence à grande échelle, la plate-forme logicielle se charge de gérer et de maintenir les composants de l'infrastructure matérielle à l'aide d'un élément fonctionnel distinct : le gestionnaire de périphériques.

Ce dernier aide les utilisateurs à configurer de manière dynamique les périphériques et les composants de l'infrastructure de visioconférence. Une fois les composants matériels déployés sur le réseau et sur l'infrastructure, le gestionnaire de périphériques surveille les problèmes pouvant survenir et aide à les résoudre. Lorsque des mises à jour logicielles sont requises, le gestionnaire de périphériques facilite leur déploiement.

Si la visioconférence connaît un tel engouement, c'est en partie grâce à la simplicité avec laquelle les appels peuvent être établis par l'utilisateur. L'intégration à Microsoft Outlook ou la création de portails de planification conviviaux facilitent désormais la planification et la gestion des appels.

Le gestionnaire de périphériques permet également de générer des rapports et fournit des détails exhaustifs sur les appels vidéo qui, une fois traités, permettent d'évaluer l'utilisation actuelle du système et les projets d'expansion du réseau vidéo.

## **11. Sécurité**

Nombre d'entreprises qui ont investi dans la visioconférence devront tôt ou tard permettre à leurs collaborateurs mobiles ou télétravailleurs de se connecter au réseau de l'entreprise afin de participer à des appels vidéo avec leurs collègues. Mais pour cela, la plate-forme logicielle doit fournir les fonctionnalités nécessaires.



Parallèlement, ces entreprises émettront également le désir de communiquer avec leurs partenaires et leurs clients à l'aide de cette technologie. Mais pour cela, le trafic vidéo doit être capable de traverser les pare-feu de façon sécurisée d'un client à un autre. Cette opération est particulièrement ardue pour la vidéo, car les pare-feu de données tentent de réorganiser les paquets de données. La mise en œuvre d'un pare-feu vidéo tel que le Polycom® VBP® (H323) résout le problème.

## **12. Gestion du contenu**

Jusqu'ici, le principal facteur qui poussait les entreprises à recourir à la visioconférence était économique, puisque la visioconférence réduit les coûts de déplacement professionnel. Depuis peu, les entreprises prennent cependant conscience que les avantages de la visioconférence peuvent s'étendre à des domaines divers et variés : formation, marketing, enseignement, conformité, communications internes, publicité, relations publiques, pour n'en citer que quelques-uns.

A mesure que l'utilisation de la visioconférence dans ces domaines gagne du terrain, les clients en découvrent tout le potentiel. En effet, cette technologie permet non seulement de communiquer en temps réel, mais aussi de revoir les enregistrements numériques d'événements et de communications passés.

Loin de se limiter aux réunions, cette même technologie peut être utilisée pour créer du contenu multimédia riche encapsulé, qu'il est possible de modifier, d'améliorer, d'archiver et de diffuser sur plusieurs supports. Ces contenus peuvent être mis à la disposition des publics cibles sur demande.

Par exemple :

- Multidiffusion d'événements en direct : la plate-forme logicielle permet de transmettre des webcasts enregistrés en streaming et prend en charge les

modes push et pull de transfert de contenu vidéo à partir des serveurs de streaming.

- Vidéo à la demande : la plate-forme logicielle automatise la création de versions archivées de n'importe quel Webcast d'événement en direct et permet ainsi aux clients de les visionner à la demande, selon leurs besoins.
- Gestion de contenu multimédia : la plate-forme logicielle permet de contrôler la manière dont les contenus vidéo sont regroupés, approuvés, classés, modifiés et publiés.
- Stockage et archivage : la plate-forme logicielle définit des règles quant à la durée du stockage du contenu vidéo gourmand en bande passante ; les clients peuvent déterminer la manière dont le contenu est conservé, transcodé et stocké dans le cloud ou sur des ressources d'entreprise, sans qu'il y ait besoin d'une intervention manuelle quotidienne.

## **Conclusion**

Dans ce chapitre nous avons présenté les normes et standards utilisés dans le domaine de visio-conférence.

La norme H.323 est le standard le plus utilisé pour la visio-conférence dans l'architecture TCP/IP, elle donne une performance de haute qualité en rapport avec le débit utilisé.

# **Chapitre II : Conception de l'application**

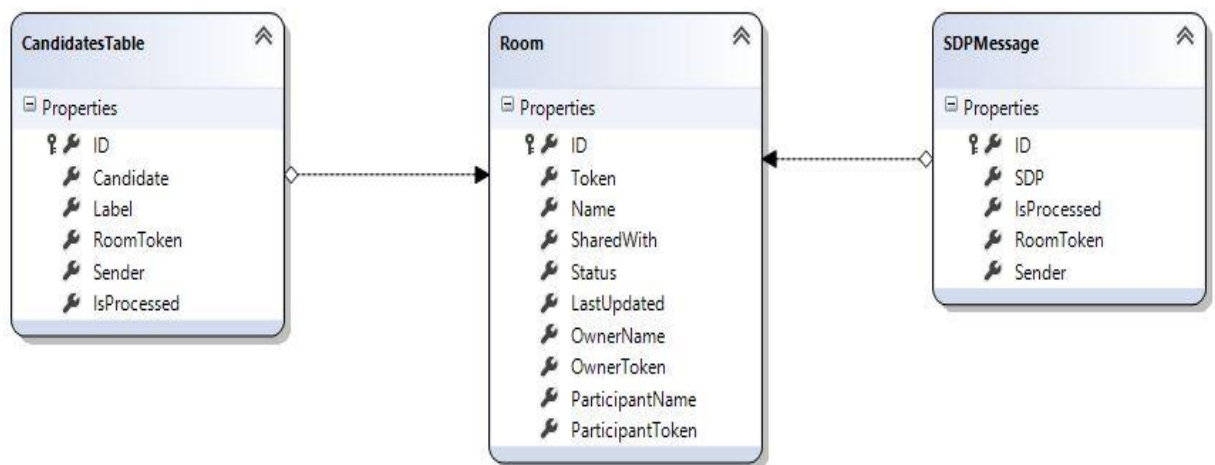
## Introduction

Dans ce chapitre nous allons réaliser la conception de l'application en utilisant le diagramme de classe et le diagramme d'activité de la méthode UML.

Ensuite nous allons présenter les outils que nous avons choisis pour codifier les flux de données audio et vidéo.

### 1. Diagramme de classe

Le diagramme se compose de trois classes :



**Figure II.1** diagramme de classe

La classe Room désigne une métaphore d'une salle de conférence, elle contient les informations d'identification de la salle, les utilisateurs participant et l'état de la salle.

La classe candidats contient les informations des utilisateurs, la classe message contient les messages de control d'accès aux salles de conférence.

La synchronisation des utilisateurs et la gestion des salles de conférence est traité pas le serveur, mais le flux multimédia passe directement en mode point a point.

## 2. Diagramme d'activité

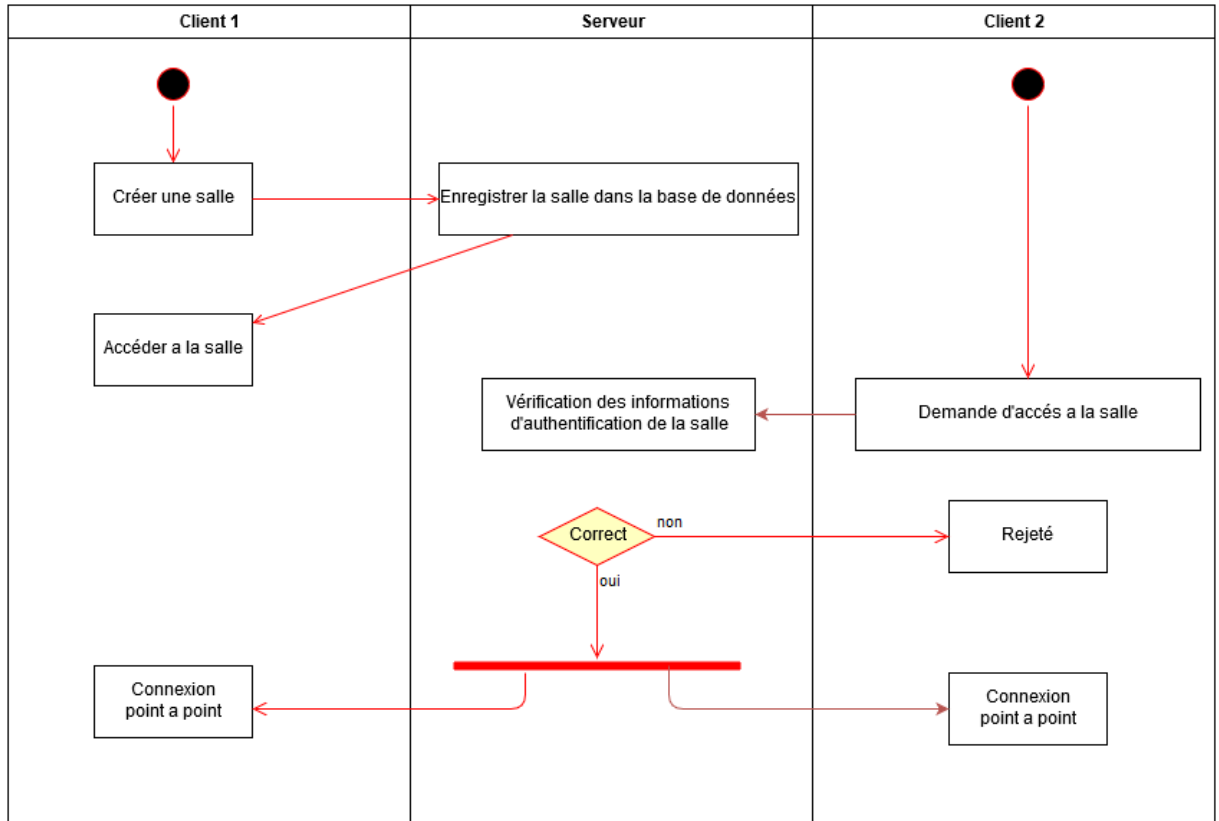


Figure II.2 diagramme d'activité

Ce diagramme présente les activités réalisées lors de l'établissement d'une session de visio-conférence.

## 3. Choix des outils

### a. Codec Vidéo

Le choix du codec vidéo est déterminé par plusieurs facteurs, les plus importants sont la disponibilité, licence, consommation de bande passante et performance. Notre choix initial a été le codec H.264, qui est pris comme la norme de codage vidéo dans l'industrie de vidéo conférence.

Les navigateurs web présents supportent le décodage H.264 nativement, ce qui veut dire qu'un lecteur HTML 5 est suffisant pour que les navigateurs soient capables de lire un flux vidéo H.264 en utilisant la balise <video>.

Mais ce n'est pas le cas pour l'encodage, comme la norme H.264 est une norme propriétaire payante, les navigateurs ne l'ont pas implémenté totalement, ce qui veut dire que le navigateur est incapable d'encoder le flux vidéo de la caméra en H.264. La solution pour ce problème est d'implémenter le codec H.264 comme plugin ou d'utiliser l'un des codecs fournis par le navigateur.

VP8 est la norme la plus proche de H.264 en termes de performance, qualité d'image, et consommation de bande passante. VP8 est une norme gratuite et pré-implémentée dans la plupart des navigateurs modernes comme Firefox, Chrome, Opera, Edge et Safari.

La version Windows de Skype 5.5 utilise le VP8 pour la communication entre 2 personnes ou en conférences multiple. Après le rachat par Microsoft le codec a été abandonné.

Proposé au W3C comme codec vidéo (au sein de WebM), pour la balise <video>, il peut donc être lu par les navigateurs prenant en charge ce format.

Jitsi peut réaliser des vidéos s'appuyant sur ce codec depuis sa version 2.0.

Le système Android lit ce codec dans le format WebM depuis sa version 2.3. La version 4.0 offre la possibilité de l'utiliser en flux de vidéo continu (*streaming*). Le Nexus 5 basé sur la version 4.4 d'Android peut encoder et décoder matériellement ce codec. La clé Chromecast peut décoder matériellement du VP8 ou du H264.

Le service de messagerie instantanée Google Hangouts utilise ce codec.

Nous avons décidé d'utiliser ce codec comme remplacement du codec H.264 jusqu'à ce qu'il soit disponible.

### **b. Codec Audio**

AAC-LC est la norme pour le codage audio dans l'industrie de visio-conférence. Il est pré-implémenté dans la plupart des navigateurs modernes.

AAC utilise un encodage dit à débit variable (en anglais *Variable Bitrate*, parfois écrit en abrégé sous la forme *VBR*), c'est-à-dire un encodage qui permet d'adapter le nombre de bits utilisés par seconde pour encoder les données audio, à la complexité du flux à un instant donné. Comme résultat, la consommation de la bande passante est inférieure aux autres codecs fournis par les navigateurs.

## **Conclusion**

Dans ce chapitre nous avons réalisé la partie conceptuelle de l'application.

Comme le résultat principal de ce travail n'est pas la conception mais l'étude de faisabilité et la réalisation de l'application, nous ne sommes pas allés au fond de la partie génie de logiciel pour réaliser cette conception.

**Chapitre III :**  
**Implémentation de**  
**l'application**



## **Introduction**

Dans ce chapitre nous allons présenter les étapes de réalisation de l'application.

Objectif de cette application est de donner à l'utilisateur la possibilité de réaliser une vidéoconférence point à point en utilisant un site web sans utiliser aucun plugin.

### **1. Langage de programmation**

Afin de réaliser notre objectif, nous avons choisis le langage ASP.NET MVC avec C#.

ASP.NET est une technique pour générer à la demande des pages web, lancée par Microsoft en juillet 2000, et utilisée pour mettre en œuvre des applications web. Il s'agit d'une évolution majeure d'Active Server Pages (ASP, alias Classic ASP), par laquelle cette technique a été incorporé dans la plateforme Microsoft .NET.

Le moteur d'ASP.NET est un filtre branché sur le serveur web Internet Information Services (IIS). Il est distribué avec le framework .NET. ASP.NET peut être utilisé avec n'importe quel langage de programmation pour la plateforme .NET (Visual Basic .NET, C#, JScript...). [14]

Depuis la version 3.5 du Framework .NET, Microsoft propose sous forme d'extensions, un nouveau modèle de conception et de développement d'applications Web, nommé ASP .NET MVC. Nous verrons qu'il ne s'agit en aucun cas d'une technologie remplaçant la technologie ASP .NET WebForms, que nous avons utilisé dans le chapitre précédent, mais d'une alternative. En effet, le modèle MVC est un modèle de développement reconnu ayant fait ses preuves dans d'autres technologies telles que les technologies J2EE et PHP. Microsoft a simplement décidé de proposer une implémentation de ce modèle, que nous allons étudier pour créer une application Web.

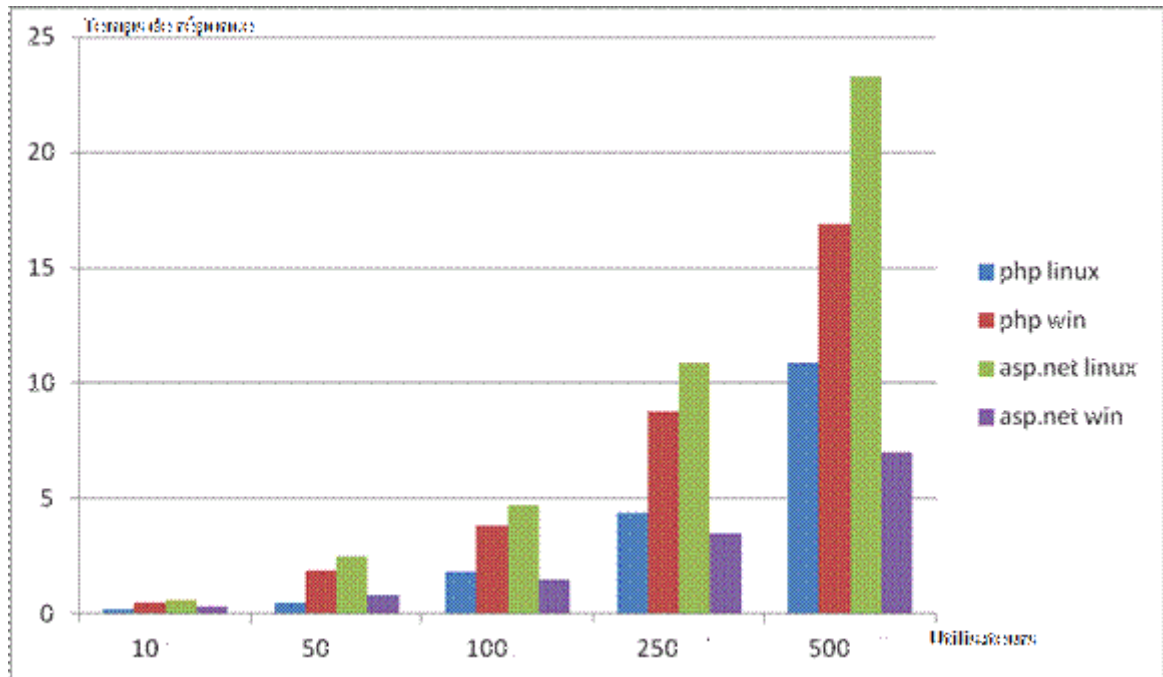
Le modèle ASP .NET MVC, où MVC est l'acronyme de **M**odèle **V**ue **C**ontrôleur, permet de créer des applications Web composée :

D'un modèle, constitué d'un ensemble de classes permettant de créer les objets métiers manipulés dans l'application, et d'exécuter les traitements métiers.

De vues constituant des éléments graphiques tels que des contrôles utilisateurs, des pages Web ou encore des Master Pages. Ces éléments graphiques sont implémentés de manière radicalement différente par rapport à leurs homologues en ASP.NET WebForms.

De contrôleurs permettant de piloter l'application, d'exécuter des actions et fournir une vue en réponse aux requêtes reçues. L'une des fonctionnalités fondamentales des contrôleurs est d'assurer la communication entre le modèle et la vue. [14]

La performance de MVC peut être mieux que celle de PHP sous les bonnes conditions, la figure suivante montre le temps de réponse d'une application qui charge des données des employés, fait le tri par nom, et les affiche, cette application a été réalisé en PHP en ASP.NET MVC sous les deux systèmes Windows et Linux, les résultats sont montré pour des nombres d'utilisateurs différents :



**Figure III.1** Comparaison de performance entre PHP et ASP.NET MVC

## 2. Transmission

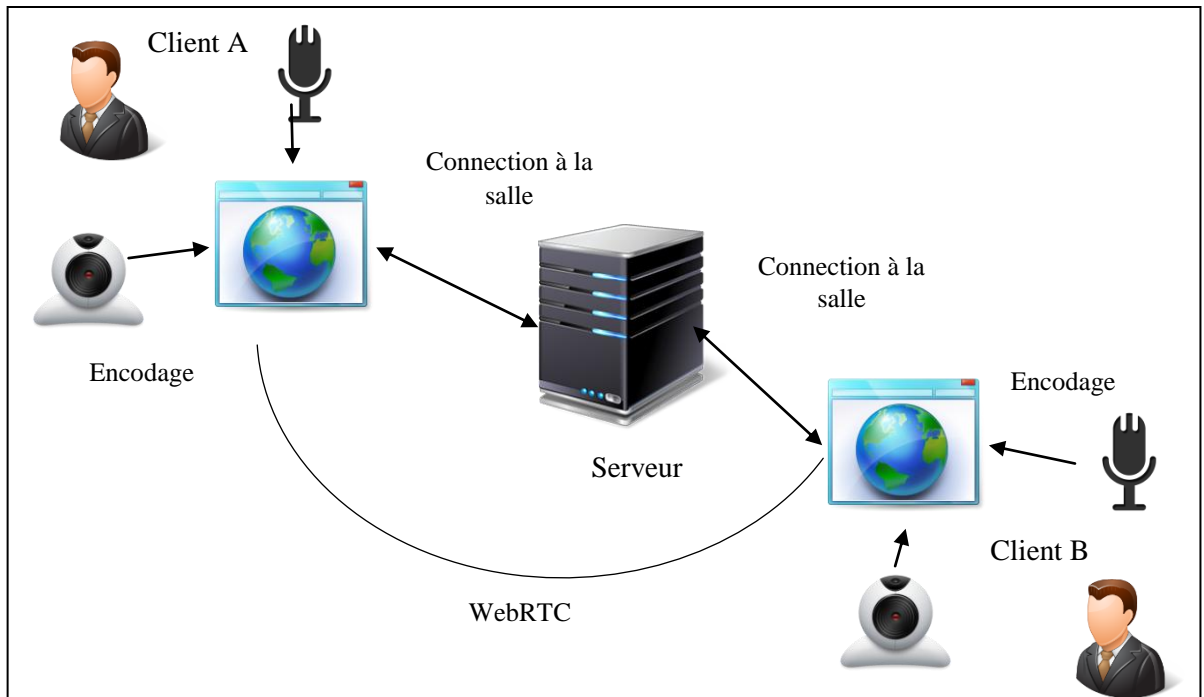
La technologie web ne supporte pas nativement les flux de données continues (data streaming), il existe plusieurs solutions de streaming à partir du serveur vers le client (ex : Microsoft Live smooth streaming), mais les solutions client vers serveur sont extrêmement rare et payant.

WebRTC (Web Real-Time Communication, littéralement « communication en temps réel pour le web ») est une interface de programmation (API) JavaScript actuellement au stade de brouillon (*Draft*) développée au sein du W3C et de l'IETF. C'est aussi un canevas logiciel avec des implémentations précoces dans différents navigateurs web pour permettre une communication en temps réel. Le but du WebRTC est de lier des applications comme la voix sur IP, le partage de fichiers en mode point-a-point.

WebRTC a fait preuve de fiabilité dans les tests que nous avons menés.

## 3. Architecture de l'application

La figure suivante montre le principe de fonctionnement de l'application.



**Figure III.2** Architecture de l'application

L'application utilise la métaphore des salles de réunion, l'utilisateur crée une salle de conférence en attente d'un partenaire, ou rejoint une salle existante créée par un partenaire.

Les salles ont deux modes de sécurité, publique ou privé. L'accès à une salle publique est libre, mais l'accès à une salle privée est protégé par un jeton, sans savoir le jeton il est impossible de rejoindre la salle.

Une demande d'autorisation est affichée à l'utilisateur pour que le navigateur peut avoir accès à la caméra et le microphone. Une fois autorisé, l'image capturée à partir de la caméra est affichée.

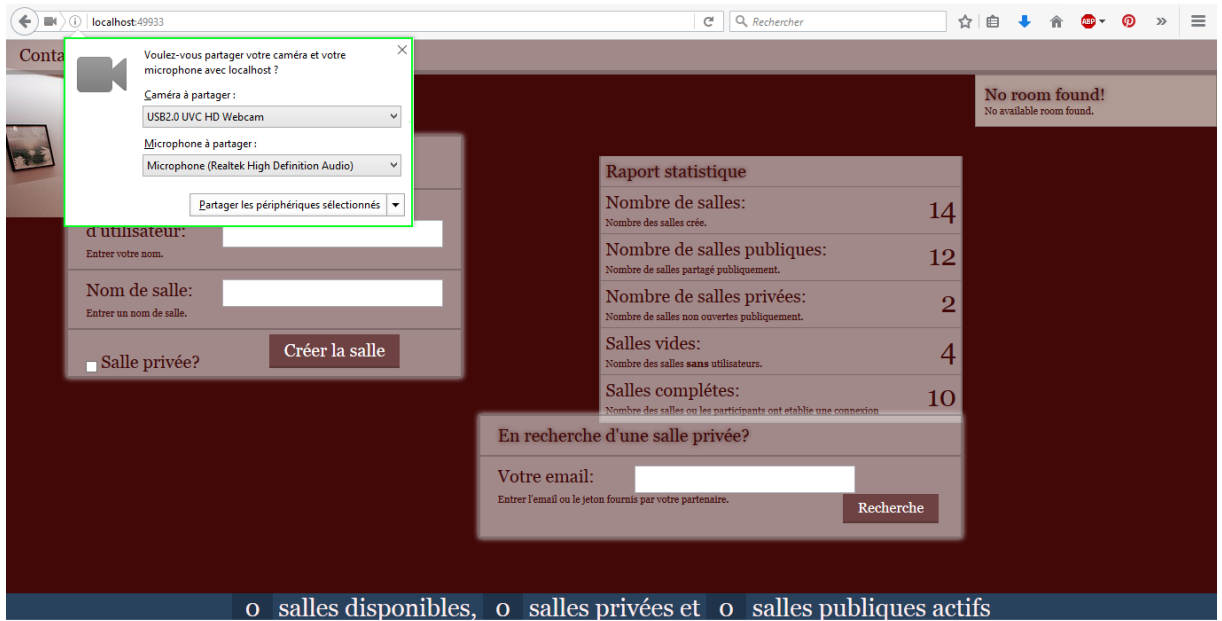


Figure III.3 Demande d'autorisation

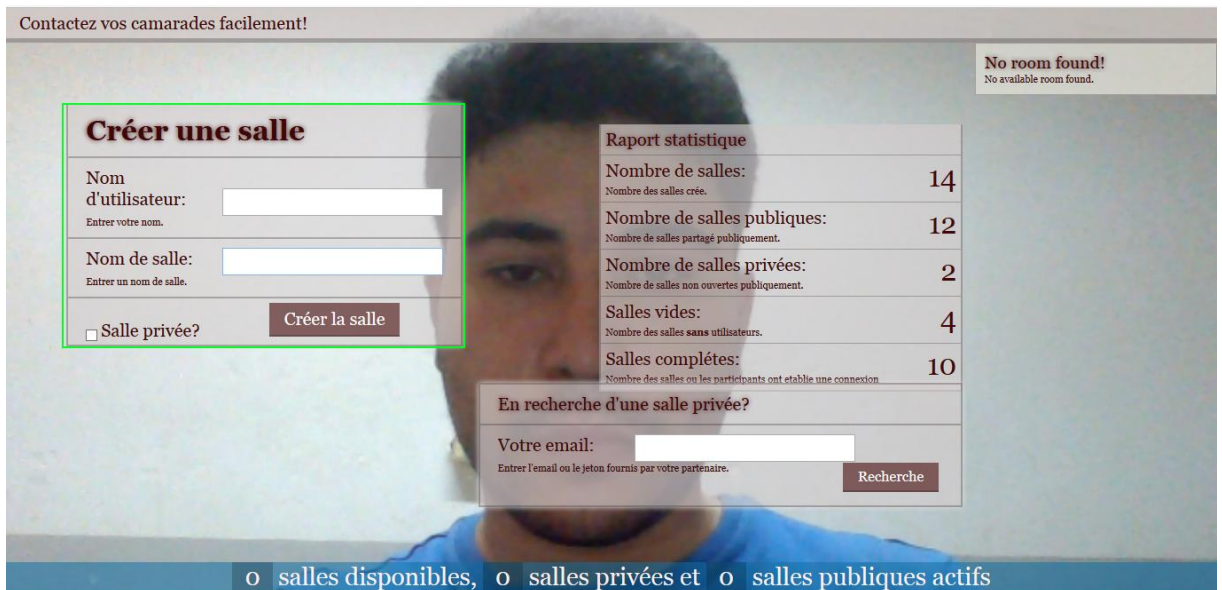


Figure III.4 La capture de la caméra est affichée

L'espace encadré représente les informations de la salle a créé, si la salle est privé, le jeton de la salle doit aussi être entré.

Pour rejoindre une salle existante, l'utilisateur peut choisir une des salles publiques comme dans la figure suivante :

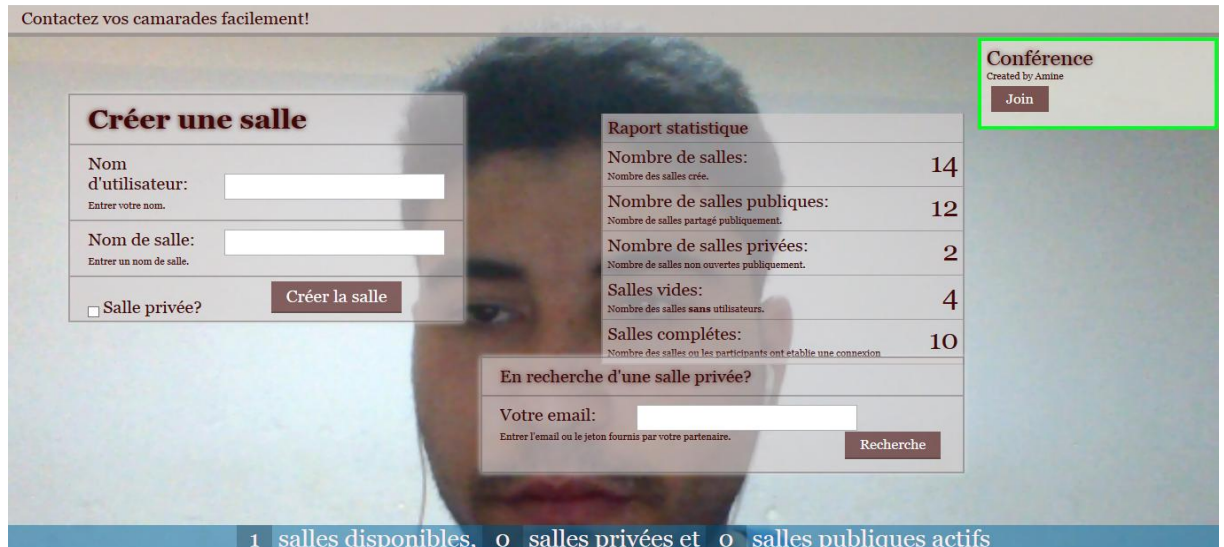


Figure III.5 Liste des salles publiques

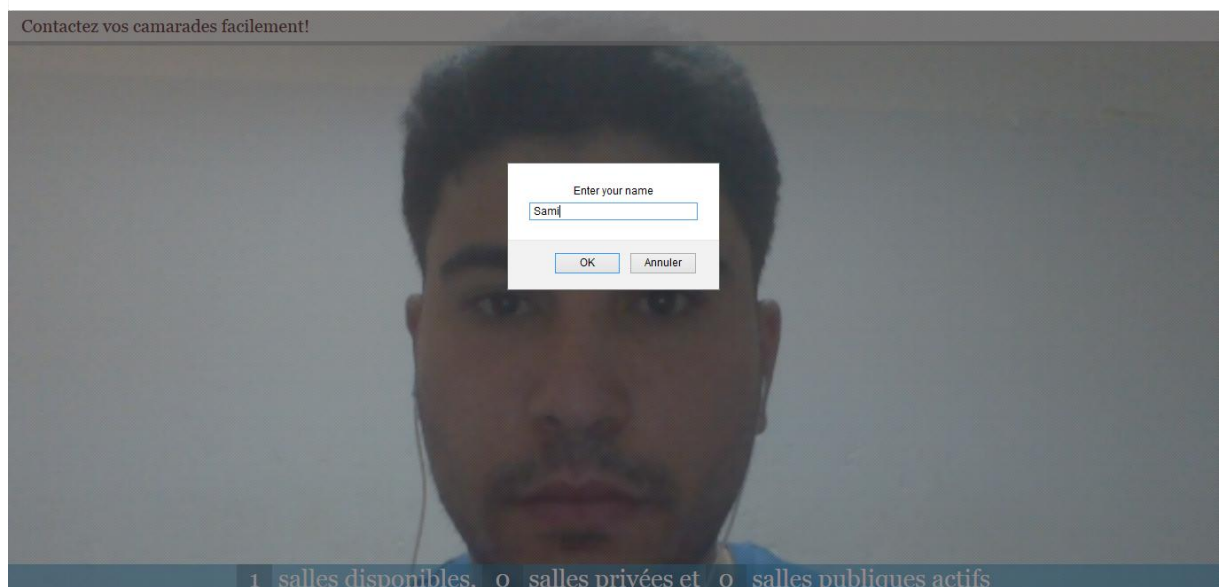


Figure III.6 L'utilisateur entre son nom pour rejoindre la salle

Si l'utilisateur cherche une salle privé, il doit entrer le jeton de la salle et lancer la recherche.

Une fois les deux utilisateurs sont connectés dans la salle, ils échangent les flux vidéo et audio entre eux :



Figure III.7 L'échange des flux Vidéo et Audio

#### 4. Performance

La figure suivante représente la consommation de la bande passante du navigateur Firefox pendant une visio-conférence avec cette application, le graph a été pris en utilisant GlassWire, use application de supervision de l'utilisation du réseau :

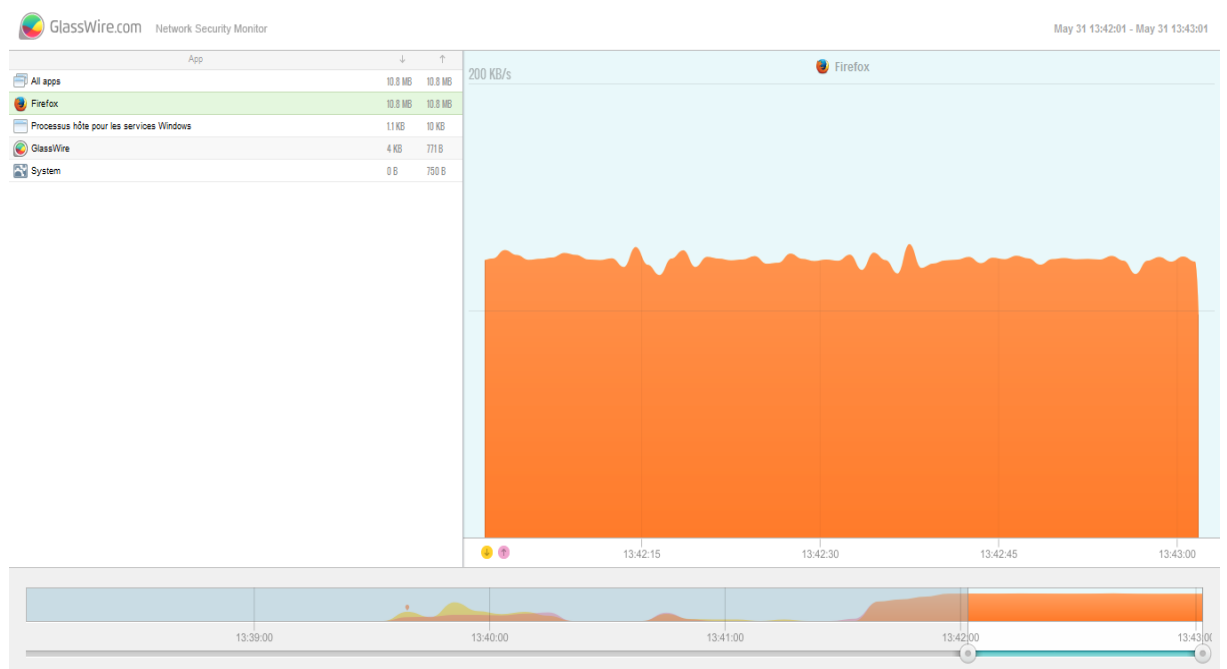


Figure III.8 Utilisation de la bande passante de l'application



Nous observons une utilisation stable de la bande passante de 170 à 190 KB/s, pour **deux utilisateurs instantanée** (85-95 KB/s pour chaque utilisateur) avec une résolution d'image par défaut de la caméra.

Pour comparer nous avons réalisé le même test avec le logiciel Skype et le site Facebook :



Figure III.9 Utilisation de la bande passante de Skype

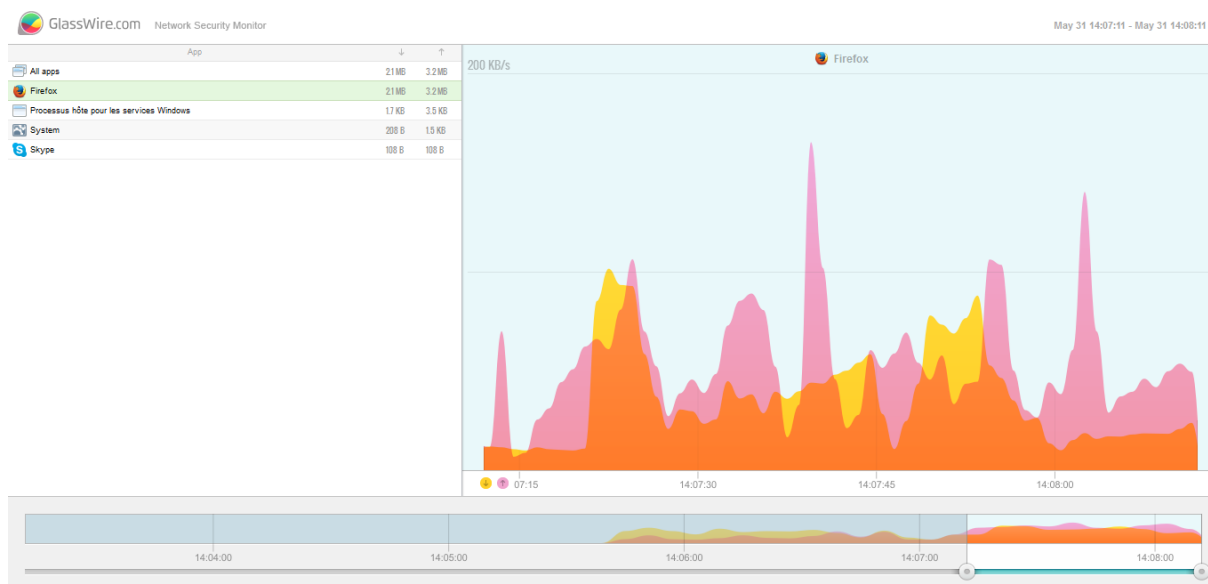


Figure III.10 Utilisation de la bande passante de Facebook



Dans le cas de Skype nous observons une utilisation de stabilité moyenne de la bande passante suivis par une qualité acceptable d'image et de son.

Pour Facebook, nous observons une grande instabilité dans l'utilisation de la bande passante et la qualité d'image et de son fournis.

Par comparaison à ces deux outils, notre application utilise plus de bande passante mais par contre fournis une qualité d'image plus élevé, une caractéristique plus demandé dans les entreprises que la faible consommation de la bande passante.

## **Conclusion**

Dans ce chapitre nous avons démontré les étapes que nous avons suivies pour mettre en œuvre cette application.

Nous estimons que 90% de notre objectif été atteint, la partie multipoint été difficile a implémenté à cause de l'architecture web qui ne prend pas en charge les flux de données continue du client au serveur.

Les tests de performance de l'application montrent une basse consommation de bande passante par rapport à la qualité de l'image en transmission.

# **Conclusion**

Face à la demande croissante pour la visioconférence, les solutions doivent fournir une infrastructure robuste, résiliente, évolutive et gérable qui offre une expérience d'une qualité incomparable.

Dans ce travail nous avons présenté les technologies modernes de la visioconférence, ainsi que notre propre implémentation d'une application web de visioconférence.

Les résultats de ce travail étés très encourageant donnant de grandes perspectives de ce domaine à cause du manque des travaux réalisé de ce type.

Les navigateurs majeurs sont en cours d'implémenter H.264, d'autre coté l'utilisation d'encodage/décodage matériel en utilisant les GPUs est capable d'introduire une nouvelle génération de logiciels de conférence.

# Bibliographies

- [01] Video Conferencing Standards and Terminology, 21<sup>st</sup> Century Video, 2015
- [02] Cisco IP Videoconferencing Solution Reference Network Design Guide, Cisco, 2002
- [03] [culture-informatique.net/comment-ca-marche-la-visioconférence/](http://culture-informatique.net/comment-ca-marche-la-visioconférence/), 05/03/2016
- [04] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), Thomas Wiegand, and all, 2002.
- [05] [fr.wikipedia.org/wiki/H.264](http://fr.wikipedia.org/wiki/H.264), 05/03/2016
- [06] [fr.wikipedia.org/wiki/MPEG-4\\_Part\\_3](http://fr.wikipedia.org/wiki/MPEG-4_Part_3)
- [07] [fr.wikipedia.org/wiki/visioconférence](http://fr.wikipedia.org/wiki/visioconférence), 01/04/2016
- [08] [fr.wikipedia.org/wiki/VP8](http://fr.wikipedia.org/wiki/VP8), 05/03/2016
- [09] [fr.wikipedia.org/wiki/webrtc](http://fr.wikipedia.org/wiki/webrtc), 10/05/2016
- [10] Introductions aux bases de la visio-conférence, Polycom, 2013
- [11] SIP : Protocole d'initialisation de session, Groupe de travail Réseau, J. Rosenberg and all, 2002.
- [12] [culture-informatique.net/comment-ca-marche-la-visioconférence/](http://culture-informatique.net/comment-ca-marche-la-visioconférence/)
- [13] [networkwebcams.com/ip-camera-learning-center/wp-content/uploads/2009/04/gop1.jpg](http://networkwebcams.com/ip-camera-learning-center/wp-content/uploads/2009/04/gop1.jpg), 15/05/2016
- [14] Développement d'une application Web avec ASP .NET MVC, James Ravaille et Julien Dollon, 2010