Communiquer différemment

http://zebra3.tel

VoIP et protocole SIP v19

1. **Principes généraux**
    1. Acoustique
    2. Traitement numérique du signal
    3. Codage vocal
    4. Evolution récente de la téléphonie
    5. Modèles réseaux
    6. Traduction
    7. Evolution des technologies $\Longrightarrow$ pourquoi SIP

2. **Le media**
    1. RTP / RTCP
    2. Dégradation de la qualité vocale
    3. DTMF
    4. Comparaison SIP – TDM

3. **Le protocole SIP (RFC3261)**
    1. Introduction
    2. Entêtes et syntaxe des messages
    3. Enregistrement des UA SIP - authentification
    4. Dialogues et transactions
    5. Adressage
    6. Media / SDP
    7. Négociation et transcodage media
    8. Cohabitation SIP / NAT

4. **Infrastructures – entités fonctionnelles SIP**

Visiter Zebra3.tel
Menu >> Formation
Télécharger le document « Cours VoIP »

1.Principes généraux

Acoustique

Définitions proposées par le dictionnaire Le Robert :

adjectif
1.1.
Qui sert à la perception des sons.
Nerf acoustique.
2.2.
Relatif au son, du domaine de l'acoustique.
Isolation acoustique.

nom féminin
1.1.
Partie de la physique qui traite des sons et des ondes sonores.
2.2.
Qualité d'un local au point de vue de la propagation du son.
Cette salle a une bonne, une mauvaise acoustique.

Acoustique : son, bruit et parole

Son : c'est la vibration d'un fluide qui se propage longitudinalement dans un milieu sous la forme d'ondes pouvant être quantifiées sur une échelle de fréquences. Il est constitué d'un mélange de plusieurs fréquences et ne se propage pas dans le vide. La mesure du niveau sonore, reliée à l'intensité d'un son exprimée en décibels (dB), correspond à l'amplitude de l'onde qui lui est associée.
La mesure du niveau sonore (pression exercée sur l'oreille par la vibration de l'air) est évaluée en décibels « pondération selon courbe A pour les faibles pressions acoustiques adaptée à l'oreille » et noté (dB (A)).

Bruit : des erreurs sont introduites lors du traitement du signal notamment en phase de quantification. Ces approximations, appelées bruit, ne peuvent plus être éliminées une fois qu'elles ont été introduites dans la chaine numérique. Le rapport signal sur bruit (SNR) est un indicateur de la qualité d'une information. C'est le rapport des puissances entre le signal d'amplitude maximale et le bruit de fond. Il s'exprime en décibel (dB).

Parole : elle est formée de sons complexes eux-mêmes constitués d'une superposition de vibrations simples d'amplitudes et fréquences différentes. Elle est principalement composée de sons voisés (vibration des cordes vocales « zzzzz »), non voisés (absence de vibration des cordes vocales « sssss »), et plosifs « ciao ». Les fréquences audibles par une oreille humaine sont situées entre 20 Hz et 20 kHz.

Gamme tempérée

- Exemple de son pur : la_3 = 440Hz donné par le diapason, fréquence de référence
- Fréquence d'une note = 440 x r^n où r = 1,05946 et ^{n} = nombre de demi-tons entre la note et le la_3
- 12 demi-tons par octave
- Monter d'une octave équivaut à multiplier la fréquence par 2
- Baisser d'une octave équivaut à diviser la fréquence par 2
- Monter d'un demi-ton équivaut à multiplier la fréquence par r
- Baisser d'un demi-ton équivaut à diviser la fréquence par r

Octave	Do	Do#/Réb	Ré	Ré#/Mib	Mi	Fa	Fa#/Solb	Sol	Sol#/Lab	La	La#/Sib	Si
9	16744.032	17739.680	18794.544	19912.112	21096.160	22350.592	23679.616	25083.712	26579.488	28160.000	29834.464	31608.512
8	8372.016	8869.840	9397.272	9956.056	10548.080	11175.296	11839.808	12541.856	13289.744	14080.000	14917.232	15804.256
7	4186.008	4434.920	4698.636	4978.028	5274.040	5587.648	5919.904	6270.928	6644.872	7040.000	7458.616	7902.128
6	2093.004	2217.460	2344.318	2489.014	2637.020	2793.824	2959.952	3135.964	3322.436	3520.000	3729.308	3951.064
5	1046.502	1108.730	1174.059	1244.507	1318.510	1396.912	1479.976	1567.982	1661.218	1760.000	1864.654	1975.532
4	523.251	554.365	587.329	622.253	659.255	698.456	739.988	783.991	830.609	880.000	932.327	987.766
3	261.625	277.182	293.664	311.126	329.627	349.228	369.994	391.995	415.304	**440.000**	466.163	493.883
2	130.812	138.591	146.832	155.563	164.813	174.614	184.997	195.997	207.652	220.000	233.081	246.941
1	65.406	69.295	73.416	77.781	82.406	87.307	92.498	97.998	103.826	110.000	116.540	123.470
0	32.703	34.647	36.708	38.890	41.203	43.653	46.249	48.999	51.913	55.000	58.270	61.735
-1	16.351	17.323	18.354	19.445	20.601	21.826	23.124	24.499	25.956	27.500	29.135	30.867

Amplitude et sa représentation

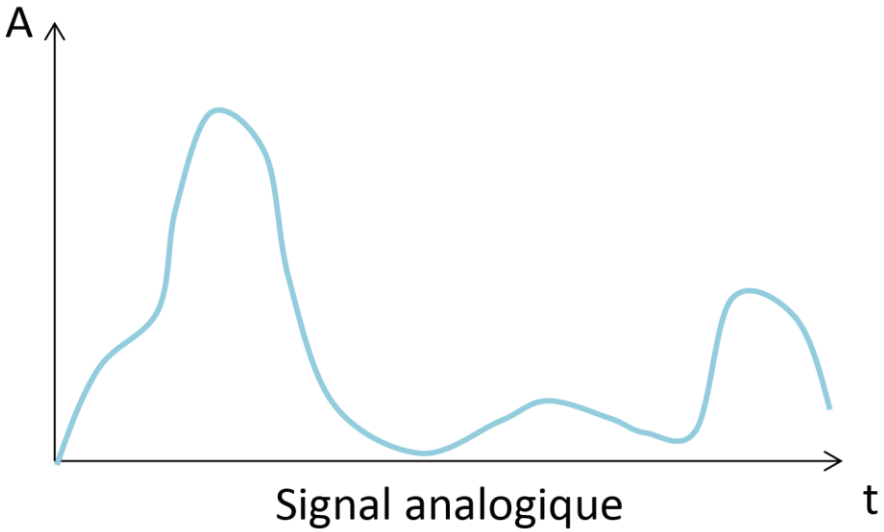
Pression (Pa)	Pression (Bar)	L_p (dB)	I (W/m^2)	Correspondance
2 000 000	20	220	10 000 000 000	
200 000	2	200	100 000 000	
101 300	1,013	194	25 118 864	Pression atmosphérique
20 000	0,2	180	1 000 000	Fusée
2 000	0,02	160	10 000	
200	0,002	140	100	Jet
20	0,0002	120	1	Seuil douleur
2	0,00002	100	0,01	Discothèque
0,2	0,000002	80	0,0001	Orchestre
0,02	0,0000002	60	0,000001	Rue, lieu public
0,002	0,00000002	40	0,00000001	Conversation normale
0,0002	0,000000002	20	0,0000000001	Chuchotement
0,00002	0,0000000002	0	0,000000000001	Seuil audition (20 µPa)
0,000002	0,00000000002	-20	0,00000000000001	

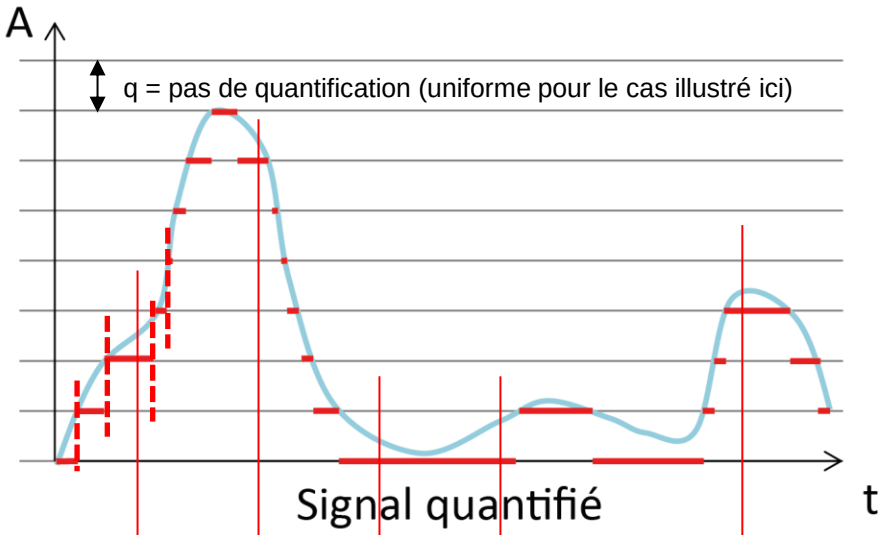
A U D I B L E

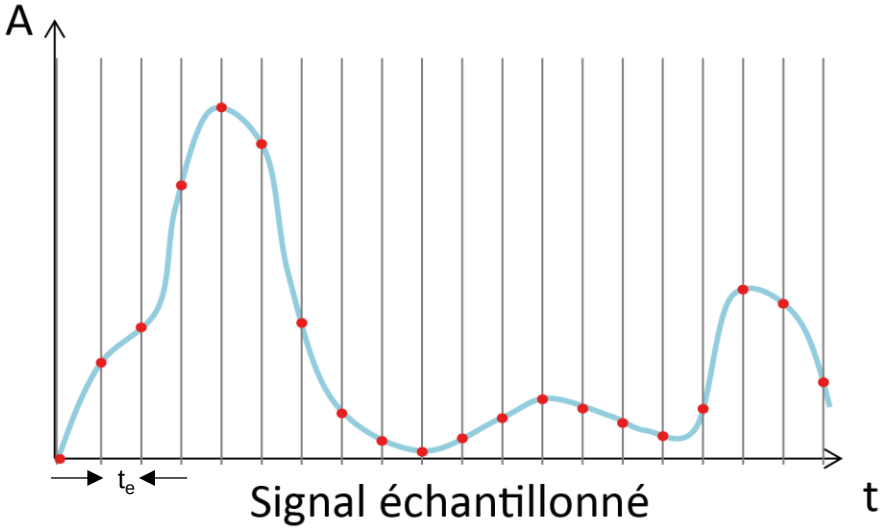
Niveau intensité sonore = $L_p = 20 \text{ Log } (P/(20.10^{-6}))$ en (dB)
Intensité sonore = $I = 10^{-12} \times 10^{L/10}$ en (W/m^2)
Puissance acoustique = P = I x S en (W)

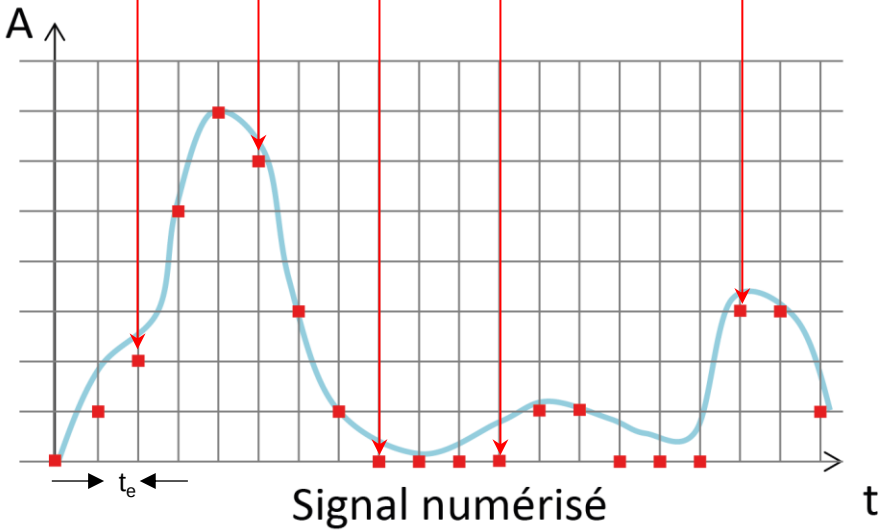
1,013 bar exerce une force de 1,032 kgf/cm^2 soit une colonne d'eau de hauteur 1,032 m
Pressions atmosphériques extrêmes relevées (hPa) : 870 – 1083

Transformations du signal





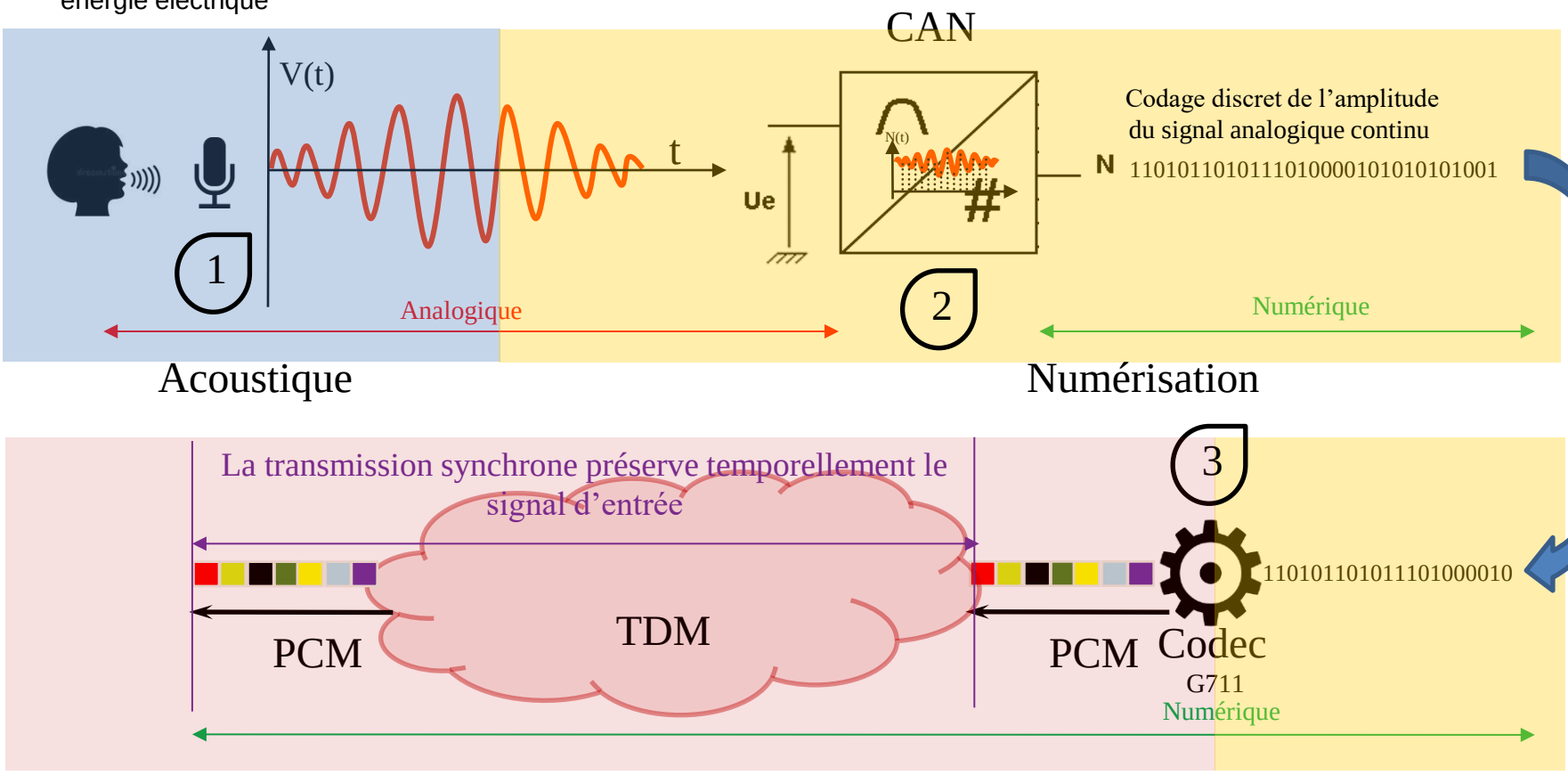




Traitement du signal voix et transport TDM

Le micro (classique) convertit l'énergie acoustique qui frappe sa membrane en énergie électrique

Le convertisseur analogique/numérique transforme la tension analogique reçue en entrée en signaux numériques. Pour la voix, la résolution des CAN est de 12 bits (quantification) /échantillon



Transmission synchrone / TDM :

1. Numérisation (et compression) de la voix / Transmission synchrone PCM codec G711 64kbps
2. Emission d'un octet (échantillons) toutes les 125 µs
3. **Multiplexage temporel** garantie de remise

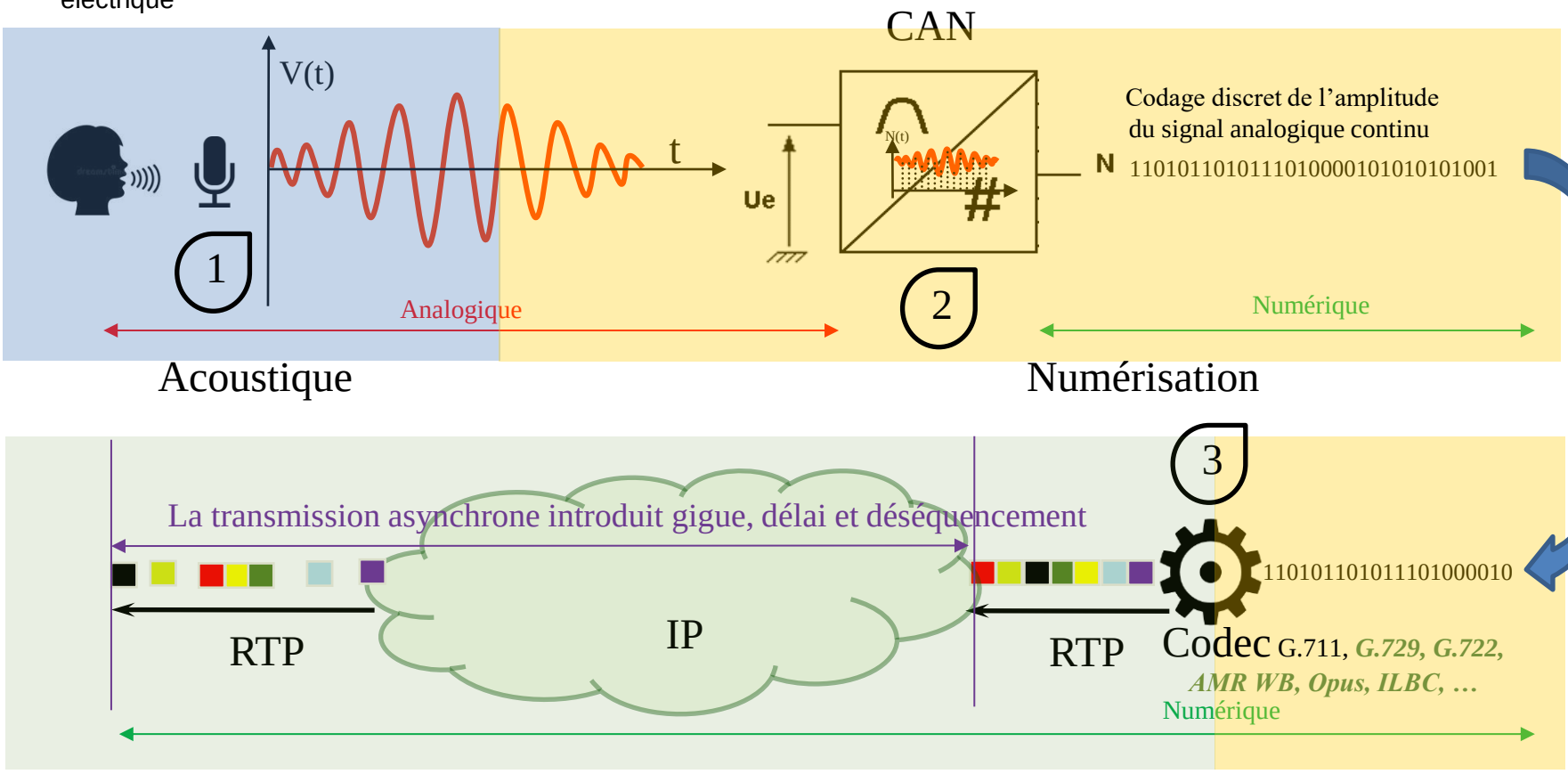
Pour numériser un signal analogique il faut :

1. L'échantillonner cad passer d'un espace temps continu à un espace temps discret
2. Le quantifier cad passer d'un espace de valeurs continu à un espace de valeur discret
3. Le coder cad valoriser numériquement chaque valeur quantifiée

Traitement du signal voix et transport IP

Le micro convertit l'énergie acoustique qui frappe sa membrane en énergie électrique

Le convertisseur analogique/numérique transforme la tension analogique reçue en entrée en signaux numériques. Pour la voix, la résolution des CAN est de 12 bits (quantification) /échantillon



Transmission asynchrone / IP :

1. Numérisation (et compression) de la voix / Transmission asynchrone par paquets RTP
2. Emission d'un paquet RTP pouvant contenir plusieurs échantillons (dépend du Codec sélectionné)
3. **Multiplexage statistique** pas de garantie de remise (drop)

Pour numériser un signal analogique il faut :

1. L'échantillonner cad passer d'un espace temps continu à un espace temps discret
2. Le quantifier cad passer d'un espace de valeurs continu à un espace de valeur discret
3. Le coder cad valoriser numériquement chaque valeur quantifiée

Impact bande passante

En VoIP, c'est le Codec G.711 (échantillonnage 8k Hz) avec 10ms de temps de paquétisation sans traitement spécifique des silences qui se rapproche le plus de la qualité observée dans les réseaux TDM.

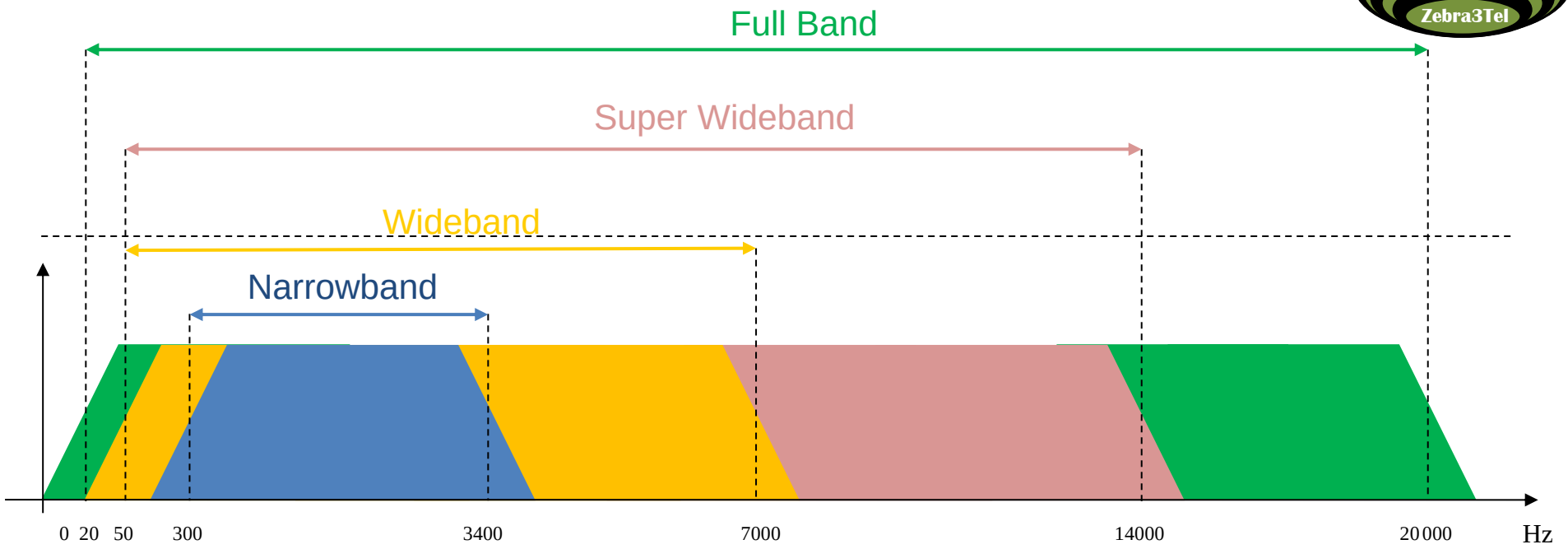
Rappel : la bande passante par intervalle de temps est : (acquisition de 8 bits toutes les 125 µs ► 8 bits / 0,000125 = 8 x 8000 = **64** kbps. Pour 10 ms acquisition de 64000 bits / 100 = 640 bits = 80 octets.

En considérant les 40 octets d'en-têtes IP, UDP et RTP pour une charge utile de 80 octets, la bande passante nécessaire pour écouler ce flux média est : 8 bits x (40 + 80 octets) x 100 = **96** kbps.

Dans ce cas de figure le plus défavorable à la VoIP, la surconsommation bande passante est de 50%.

Il existe cependant des possibilités pour la réduire : usage de codeurs moins gourmands, augmentation le temps de paquétisation ou activation de la VAD pour réduire le volume du flux média lorsque une absence de parole est détectée. Ces alternatives apportent cependant leur lot d'inconvénients.

Spectre vocal et codeurs vocaux appliqués à la téléphonie



3 technologies de codeurs vocaux permettent de coder plus ou moins largement le spectre vocal :

1. La bande passante traditionnelle utilisée en téléphonie fixe, appelée bande-étroite ou narrowband, est 300-3400Hz. (G.711, G.729, G.723, G.726)
2. La technologie VoIP et l'augmentation des débits disponibles dans les réseaux ont permis l'émergence de codecs bande large ou Wideband plus évolués couvrant un spectre plus étendu afin d'améliorer sensiblement la qualité des conversations téléphoniques. (AMR WB, G.722)
3. Une nouvelle génération de codecs Super Wideband apporte une qualité vocale encore meilleure en couvrant le spectre 20-14000Hz. (Opus, Silk, EVS)

Codeurs vocaux

Le codec (**co**der **dec**oder) est un dispositif dont l'objectif est l'encodage ou le décodage d'un flux de données numériques afin de le transmettre ou le stocker.
Le codec réduit généralement le débit des transmissions en exploitant les propriétés de la production vocale, il introduit inévitablement des dégradations plus ou moins importantes en fonction de sa qualité intrinsèque.

Il existe 2 grandes catégories de codeurs :

1. les codeurs temporels ou à forme d'ondes PCM, ADPCM,
2. les codeurs par analyse et synthèse ABS et à prédiction linéaire CELP,

Différents algorithmes sont mis en œuvre dans la plupart des codecs :

1. quantification adaptative,
2. quantification différentielle et prédictive,
3. prédiction linéaire de signal,
4. quantification vectorielle,
5. codage entropique,

Certains codecs implémentent des traitements spécifiques pour réduire la bande passante lorsqu'il n'y a pas d'information de parole à émettre :

1. VAD pour détecter les périodes de parole (35%) et silence (65%),
2. DTX pour transmettre le minimum d'informations pendant les périodes de silence,
3. CNG pour générer un bruit de confort pendant les périodes de silence,

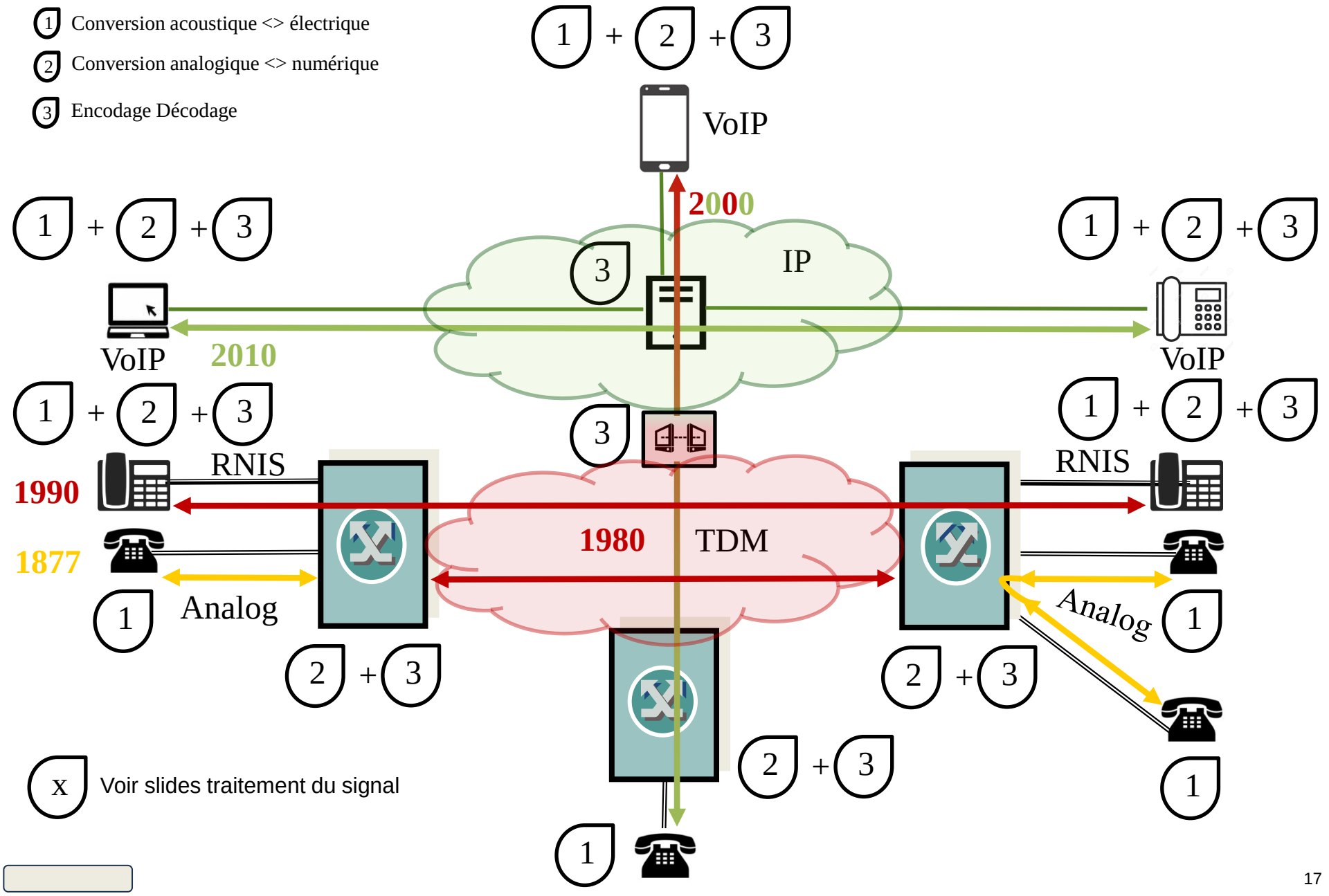
Codeurs audio les plus fréquemment utilisés

Codec	Standard	Description	Débit (kb/s)	Echantillonnage (kHz)	Remarques	MOS (Mean Opinion Score)
G.711	ITU-T	Pulse code modulation (PCM)	64	8	U-law (US, Japan) and A-law (RoW) PCM	4.1
G.722	ITU-T	7 kHz audio-coding within 64 kbit/s	64	16	SB-ADPCM (Sub-Band ADPCM)	
AMR-NB	3GPP/ETSI	Coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss	4.75/ 5.15/ 5.90/ 6.70/ 7.40/ 7.95/ 10.20/ 12.20	8	ACELP (Algebraic CELP)	
AMR-WB G.722.2	3GPP/ITU-T	Adaptive Multi-Rate Wideband Codec (AMR-WB)	23.85/ 23.05/ 19.85/ 18.25/ 15.85/ 14.25/ 12.65/ 8.85/ 6.6	16	ACELP (Algebraic CELP)	4+
G.726	ITU-T	40, 32, 24, 16 kbit/s adaptive differential pulse code modulation (ADPCM)	16/24/32/40	8	ADPCM (Adaptative PCM)	3.85
G.728	ITU-T	Coding of speech at 16 kbit/s using low-delay code excited linear prediction	16	8	LD-CELP (Low Delay CELP)	3.61
G.729	ITU-T	Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)	8	8	CS-ACELP (Conjugate Structure Algebraic CELP)	3.92
Speex	Xiph.org		8, 16, 32	2.15-24.6 (NB) 4-44.2 (WB)	CELP	
iLBC	IETF		8	13.3		
SILK	Skype		From 6 to 40	16/ 16/ 24		
OPUS	IETF	8–48 kHz LP, MDCT	From 6 to 510	Variable	Based on Skype and CELT from Xiph.org	4+

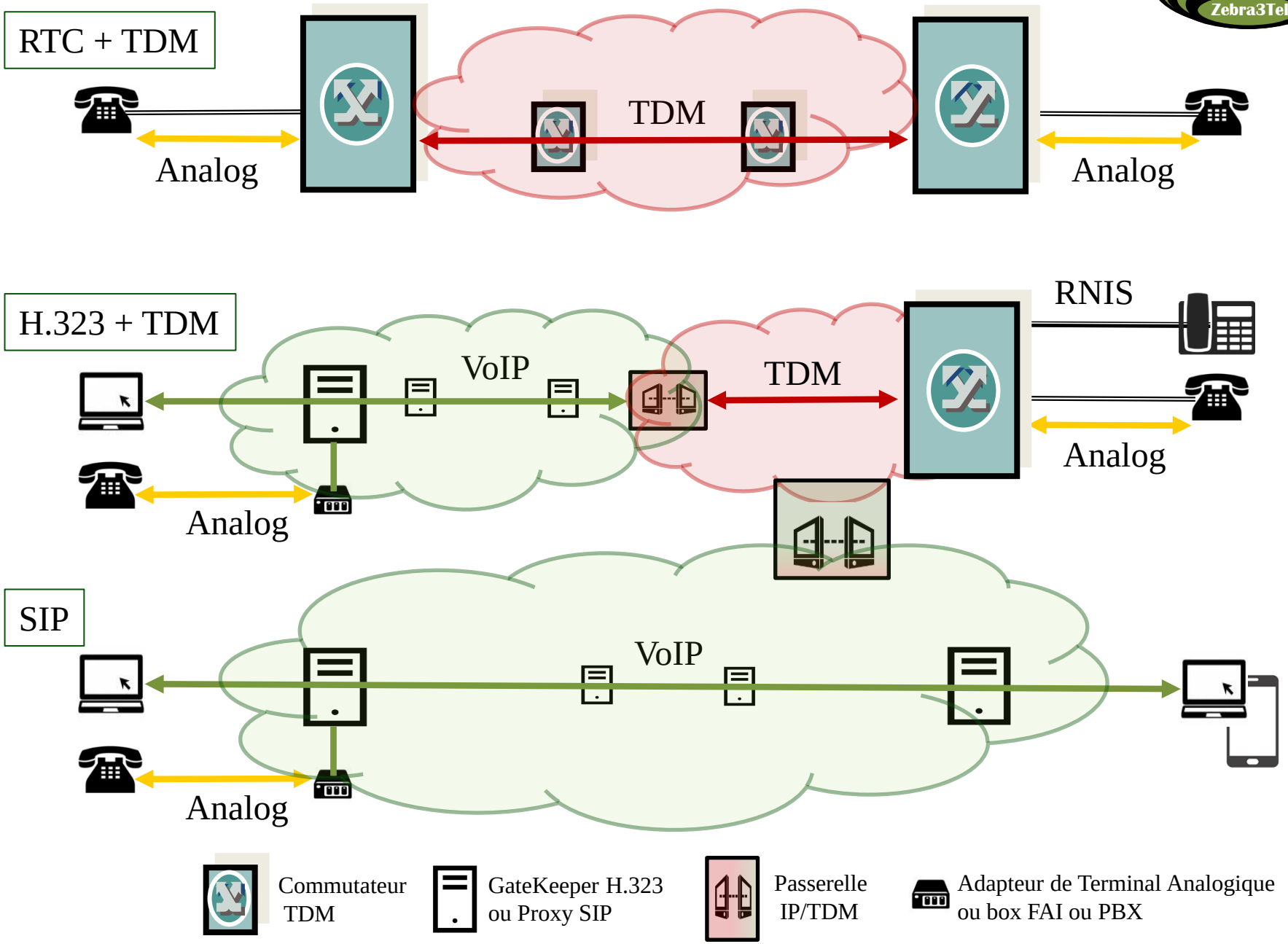
Qualité subjective des codeurs : comparatif

Contexte	Qualité	Codec	MOS P.863	Ecoute
RNIS (ISDN)	NB	G.711	2.5	
Appel 3G	WB	AMR-WB 12.65	3.5	
Référence	SWB	EVS OPUS	4.8	

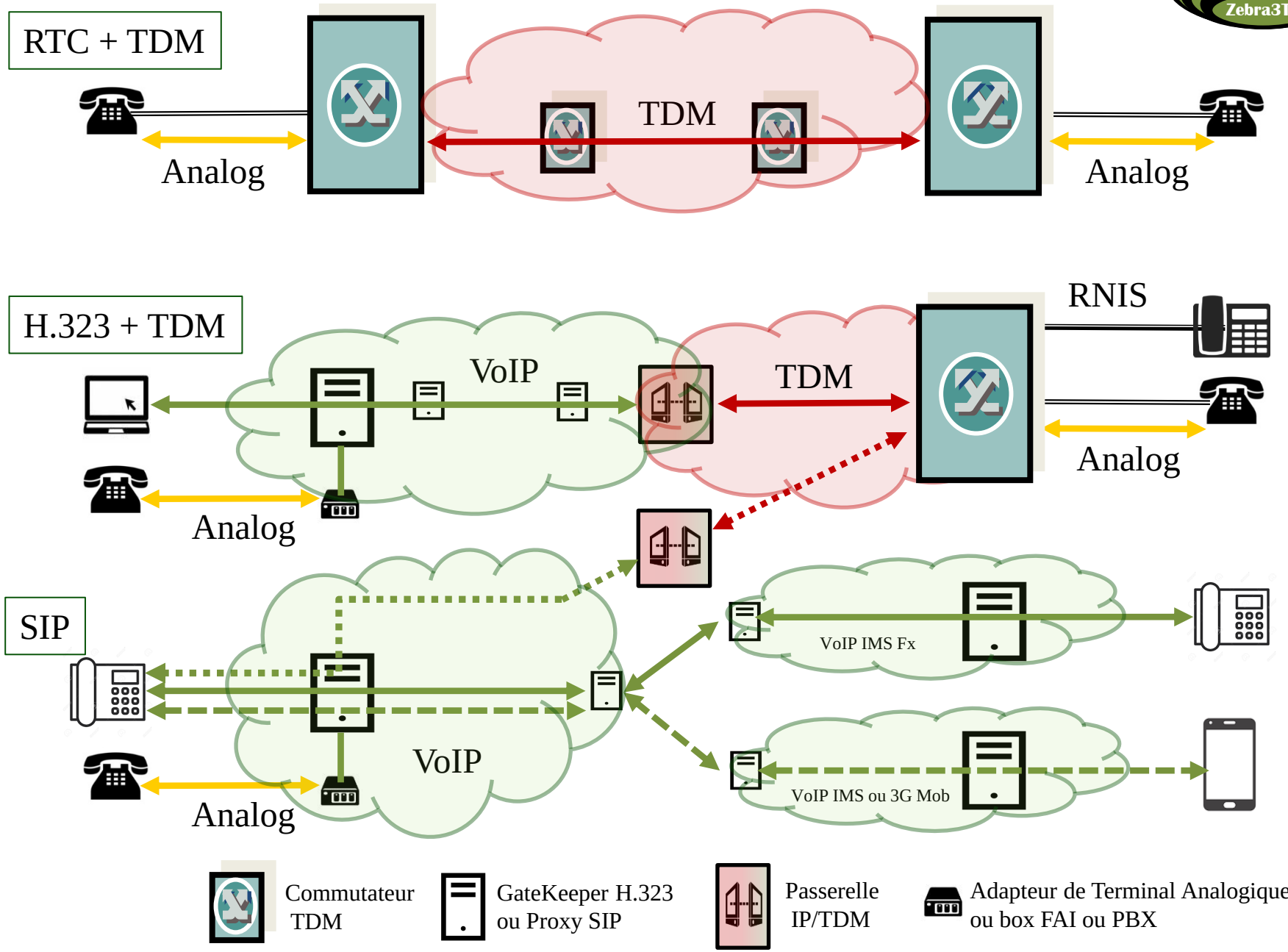
Evolution des supports utilisés en téléphonie



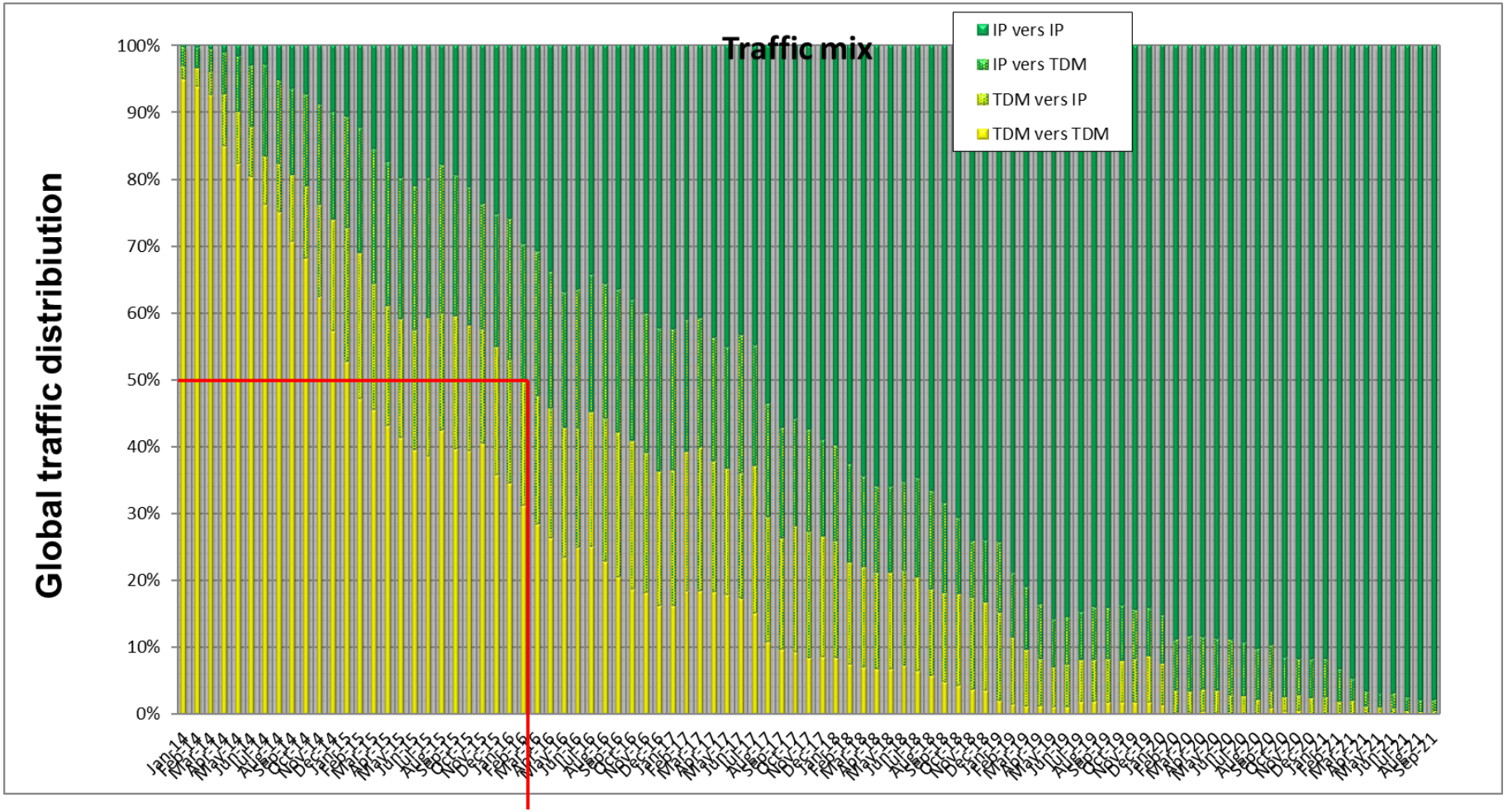
Cas d'usage les plus fréquemment rencontrés



Cas d'usage les plus fréquemment rencontrés



Evolution de la répartition des trafics TDM / VoIP



En Février 2016, les trafic TDM et VoIP véhiculés sur les réseaux de cet opérateur télécom étaient équivalents. En mai 2020 le trafic reçu en TDM était quasiment nul.

Modèle réseau maillé type P2P

Les réseaux complètement maillés imposent un nombre de liaisons élevé entre les points qui y sont raccordés. Les algorithmes de recherche du chemin optimal entre 2 points sont assez compliqués à mettre en œuvre.

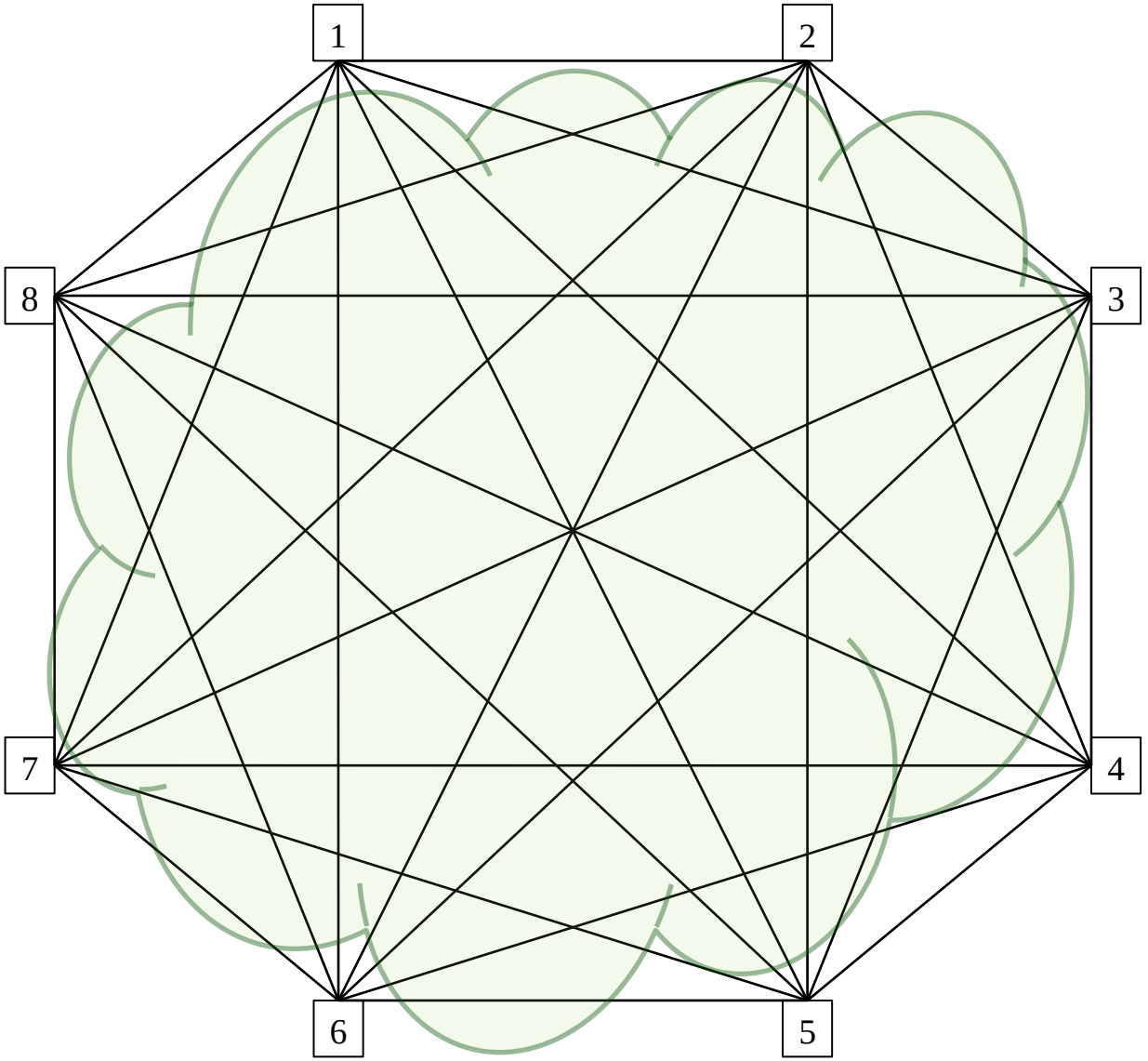
Le nombre de possibilités pour établir une communication entre 2 points parmi 8 est :

A_8^2= 8! / (8-2)! = 56.

Le nombre de liens pour mettre en relation 2 points parmi 8 est :

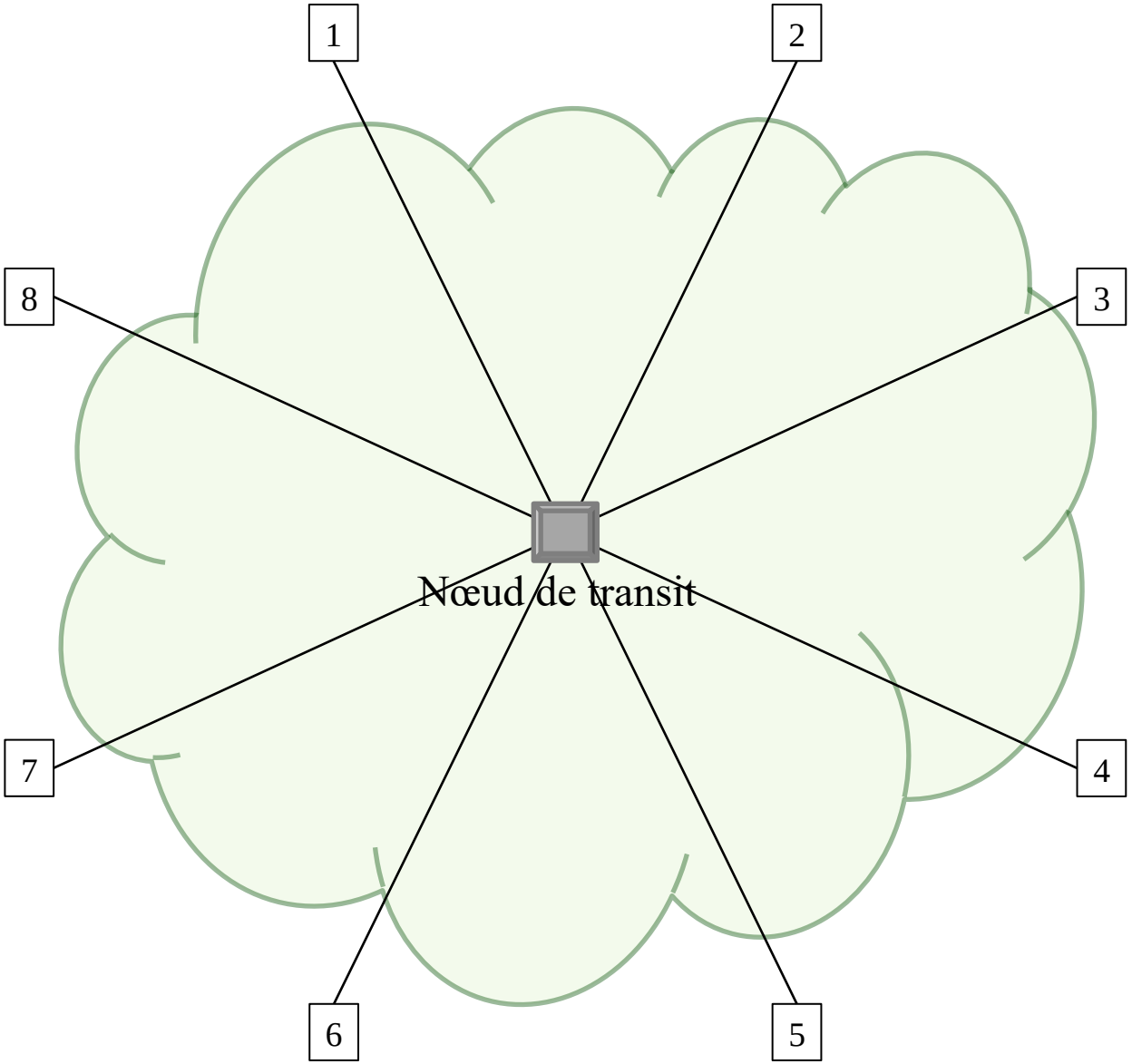
C_8^2= 8! / (8-2)! x 2! = 28.

Une infrastructure de type réseau local pourrait éventuellement convenir mais n'est pas adaptée à la téléphonie « classique » et ne permet pas, par définition de connecter des éléments géographiquement éloignés les uns des autres.

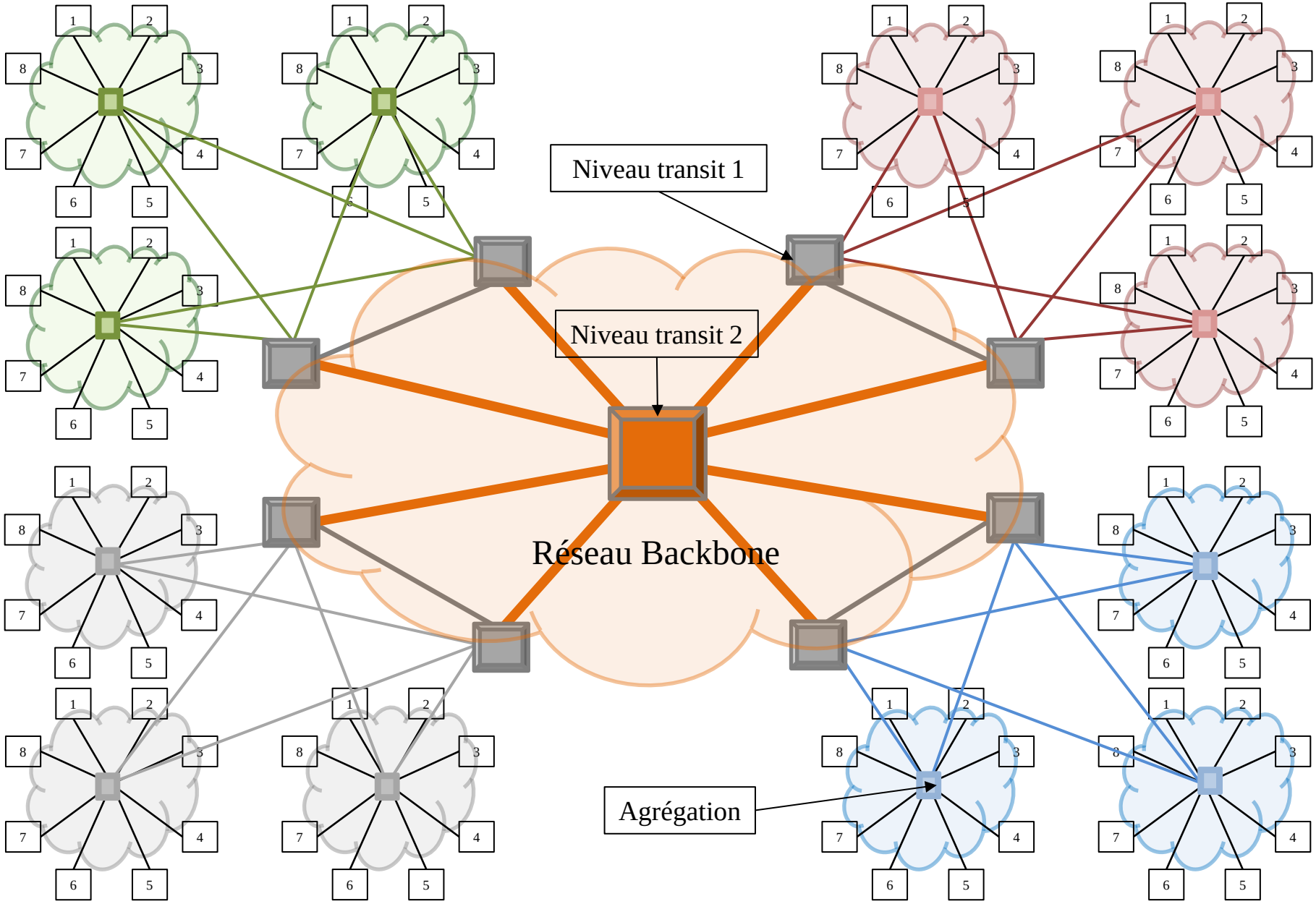


Modèle centralisé notion de transit

Il n'est pas raisonnable de concevoir une infrastructure où tous les points seraient totalement maillés.
Les architectures réseaux sont hiérarchisées et structurées en « étoile » afin de limiter le nombre d'interconnexions et d'améliorer le rendement des liens entre nœuds. Une communication entre 2 points transitera donc par un ou plusieurs nœuds intermédiaires dans lesquels seront mis en œuvre les **mécanismes de routage**. Ceux-ci permettent d'acheminer une demande d'appel après analyse du numéro demandé. L'élément central permet la fourniture de services plus ou moins évolués. Dans cet exemple 8 liaisons suffisent pour mettre en relation tous les points du modèle.



Modèle réseau hiérarchisé



Les principes de traduction

Les réseaux hiérarchisés nécessitent la mise en place de la notion de traduction permettant :

- De mécaniser la mise en relation de 2 interlocuteurs A et B
- D'établir et gérer efficacement les connexions possibles A_n^2= n!/(n−2)! entre *2* locuteurs parmi *n* abonnés
- De déterminer la meilleure voie de sortie d'un élément réseau pour faire progresser un appel dans le réseau (le terme « meilleure » reste subjectif)
- D'offrir des services à valeur ajoutées en dirigeant certains appels vers des plateformes de service spécialisée (messagerie vocale, numéros courts, portabilité, serveurs de redirection, numéros 800, …)

Les mécanismes de traduction sont mis en œuvre à chaque traversée d'un élément réseau traitant la signalisation téléphonique (commutateur d'abonné ou de transit, MSC, PBX, P et I-CSCF, SBC, IPPBX, …) quel que soit le protocole (SS7, RNIS, SIP, H.323, …)

Les différentes phases de la traduction

1. Analyse : action de confronter le numéro appelé reçu au plan de numérotage implémenté dans l'élément réseau,

2. Routage : action de déterminer un sous ensemble de routes pour aiguiller l'appel vers sa destination selon une logique propre à l'élément réseau,

3. Acheminement : sélection de la route en fonction d'un certain nombre de critères (débordement en cas de saturation de la route nominale, partage de charge, routage en fonction de l'heure ou du jour, …). Un acheminement emprunte un faisceau (trunk en SIP) pour atteindre le prochain nœud (peer SIP) jusqu'à ce que l'appel atteigne la plateforme (AS, IPPBX) qui héberge l'abonné. Pour atteindre la destination finale, la demande d'appel traverse 2 nœuds d'abonnés (class 5) et optionnellement plusieurs nœuds de transit (class 4).

On verra plus loin que c'est essentiellement la partie user de la Request-URI qui est traitée dans le cadre global de la traduction.

Eléments de contexte

- C’est le numéro appelé qui est généralement exploité lors des analyses. Le format de la numérotation téléphonique publique est défini par la recommandation itu-t E.164. Les opérateurs publics nationaux garantissent l’unicité des numéros qu’ils attribuent à leurs abonnés. Dans le cas d’installations privées, le format de la numérotation est libre, de ce fait il est très difficile d’interconnecter des réseaux téléphoniques privés hétérogènes directement entre eux tout en conservant le plan de numérotage privé.

- D’autres champs de la demande d’appel entrante ou bien des paramètres d’environnement pourraient être utilisés pour enrichir les décisions prises au niveau du routage. (exemple : origine de l’appel, identifiant de l’appelant, jour de la semaine, …). Si le numéro reçu n’est pas trouvé dans le plan de numérotage, l’appel est immédiatement libéré car il sera impossible d’associer une voie de sortie à une destination inconnue.

- Contrairement aux réseaux IP, il n’existe pas de protocole de routage dynamique (de type BGP, OSPF, RIP) pour automatiser le routage automatiser le routage des appels téléphoniques. Toute les prises de décision de routage sont locales à l’équipement traitant l’appel. Ces décisions ne sont pas forcément influencées par les autres éléments réseaux avoisinant. Cependant, des outils externes optionnels peuvent contribuer à apporter plus de flexibilité en modifiant par exemple les routes en fonction de différents critères.

De l’analyse à l’acheminement

1. Tout ou partie du numéro appelé reçu dans la signalisation entrante est confronté au plan de numérotage implémenté dans la machine. Les entrées du plan de numérotage sont constituées de préfixes qui caractérisent les destinations. Lorsqu’une correspondance suffisamment pertinente est trouvée, la phase d’**analyse** s’interrompt et un label vers un premier niveau de routage est déterminé.

2. Le label de routage de premier niveau correspond à une liste de voies de sortie (routes) vers les équipements réseau avals qui permettront de faire progresser l’appel vers sa destination. Cette liste est **constituée de routes individuelles** qui sont ordonnées selon des logiques particulières. Exemples de logique : route simple, répartition de charge plus ou moins équilibrée sur différentes routes, routes nominales avec voies de débordement, coût, …

3. La sélection définitive de la voie à emprunter dépend de l’état des routes au moment de les utiliser et de la logique décidée à l’étape de routage. Celle-ci est censée retourner la liste des meilleures routes à emprunter pour **acheminer** l’appel vers sa destination (notion de meilleur choix de routage possible).

Evolution récente des technologies

	Téléphonie		Protocoles	Data Mobile	Equipements	Systèmes OS	Internet	Transmission
	Technologies	Signalisation	Data					
Avant	Centraux electromécaniques rotatif	Analogique impulsions						CSM
1955	Centraux electromécaniques crossbar				Mainframe	IBM OS/3x0		FH
1960								Modem
1965		Analogique implusion et DTMF						Satellite
1970	Commutation électronique Spatiale		SNA Decnet		Mini informatique	RT11 / VMS		FO
1975	Commution électronique Temporelle	SS7						PCM
1980			X.25		Minitel	Unix Logiciel libre		Ethernet
1985	Radiocom2000/NMT				Micro informatique PC Macintosh	MS DOS	HTML	LAN
1990	PABX	ISDN	ATM			Windows Linux	Web	
1995	GSM					Cisco IOS	Yahoo I.E. email	
2000	VoIP	H.323	IP	WAP / GPRS		MacOS	Google	
2005	3G			UMTS	Smarphone	iOS Android	Web 2.0	ADSL
2010		SIP		Lora			Apps mobile	WiFi
2015	4G			LTE	Virtualisation	Cloud Computing	OTT	FTTH
2020	Pre 5G						IoT ou Web 3.0	

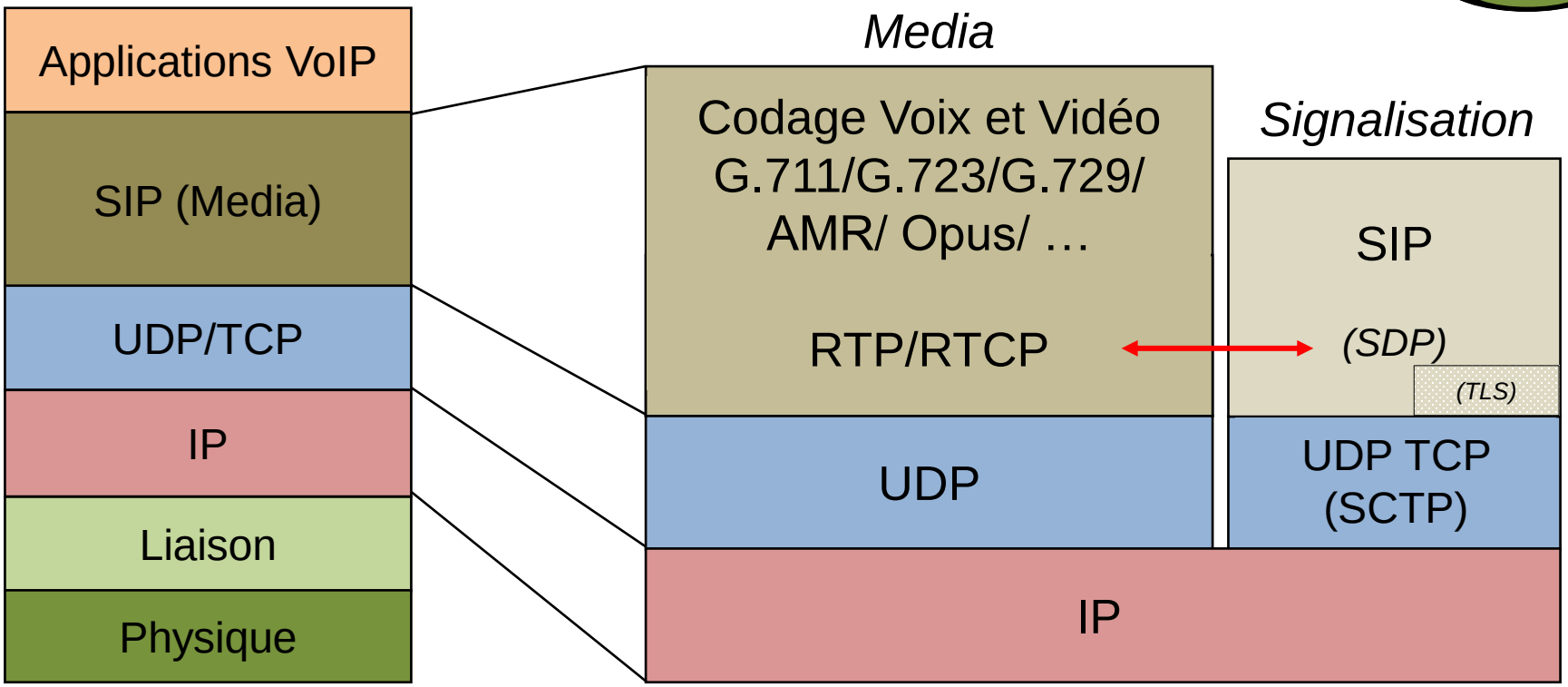
Pourquoi SIP s'impose au détriment des autres protocoles

- Call flow relativement simple, 5 messages (mini) pour établir une communication,
- Orientation multimédia même si la voix et la vidéo restent les principaux services,
- Adopté par le 3GPP pour la 4G et 5G,
- Implémenté par tous les acteurs majeurs telecom (opérateurs, terminaux, infrastructure réseau),
- Interopérabilité TDM moins nécessaire au fil du temps,
- Implémenté aussi bien dans les nœuds d'accès (class 4) que de transit (class 5),
- Tous les IP PBX open source ou constructeur sont compatibles SIP,
- Offre terminaux, applis et softphone SIP abondante,
- Fonctionne sur UDP ou TCP,
- Sécurisation TLS possible (SIPS port par défaut 5061),
- Ouverture du protocole lui permettant d'évoluer, communauté importante,
- Infrastructure virtualisable … non sans poser des problèmes de performance au niveau de l'adaptation media,
- **SIP de bout en bout = moins d'interfonctionnements de protocole, pas d'encapsulation,**

SIP n'est pas le seul protocole VoIP. De nombreuses applications VoIP comme Skype ou WhatsApp utilisent soit leur propre protocole soit l'extension JINGLE pour gérer les sessions multimedia entre entités XMPP. Asterisk propose son protocole propriétaire IAX. Ils implémentent du chiffrage spécifique. Ce « protectionnisme » rend très difficile l'interfonctionnement de ces applications avec les réseaux SIP plus ouverts et TDM.

2. Le media

Protocoles mis en œuvre par la VoIP SIP



La négociation du flux media entre 2 points s'effectue en traitant les contenus des champs des SDP échangés pendant la phase d'établissement de la communication SIP. Ce dialogue permettra de déterminer les adresses IP, ports et codec nécessaires à l'établissement de la session media. C'est la valeur du profil media retenue au niveau signalisation qui figurera dans le champ « payload type » (PT) des paquets RTP.

RTP et RTCP

- Le Real Time Protocole (RTP) RFC3550 est utilisé pour véhiculer des données - ports pairs - qui empruntent des réseaux ne garantissant pas :
  - ✓ la remise des paquets dans l'ordre où la source les a émis,
  - ✓ la remise de tous les paquets que la source a émis,
  - ✓ un délai constant entre 2 paquets consécutifs,

  La voix est un exemple de flux isochrone, c'est-à-dire qui ne tolère pas les inconvénients ci-dessus énumérés. RTP est généralement mis en œuvre au-dessus d'UDP qui garantit l'intégrité des données grâce au « checksum » et implémente la notion de port. C'est à RTP qu'incombe la gestion du séquencement des paquets d'une session. (Session = l'ensemble de tous les participants qui échangent des paquets RTP entre eux).

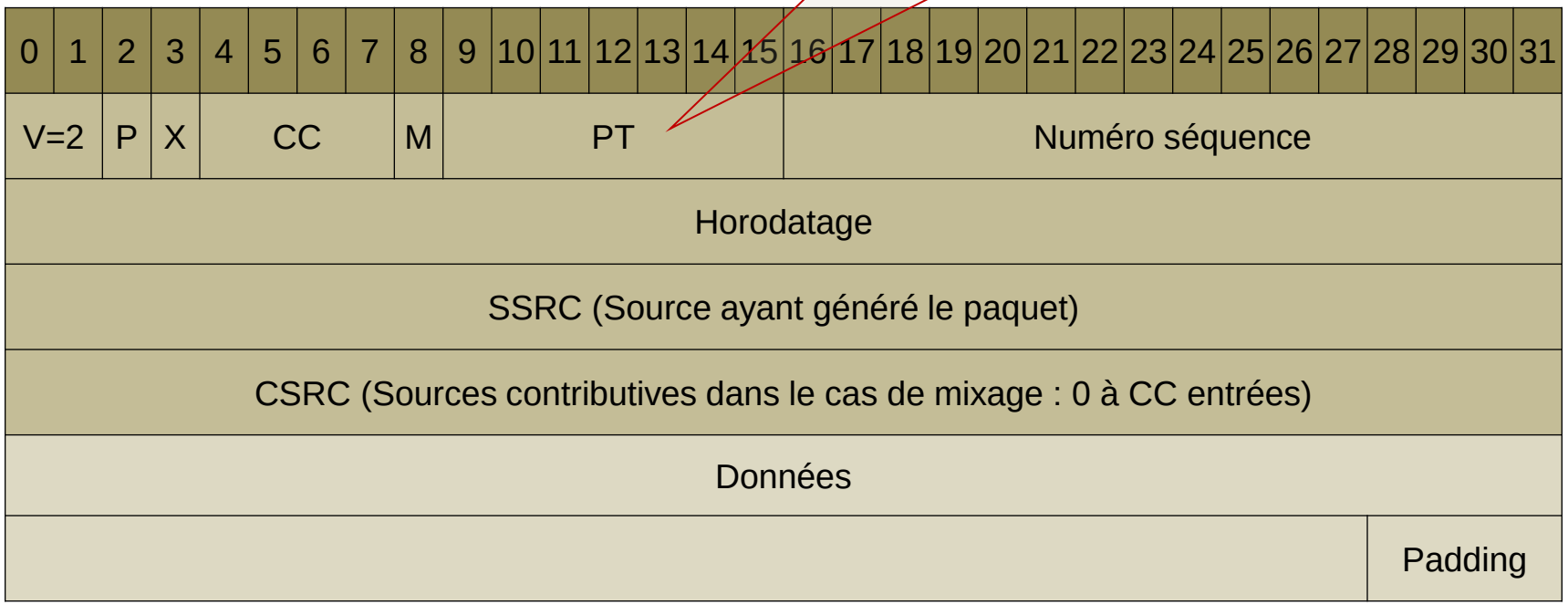
- Le Real Time Control Protocole (RTCP) - ports impairs - est utilisé pour que les participants d'une session transmettent périodiquement des informations (qualité de transmission entre autres) appelées rapports. Ainsi, l'émetteur de données recevra un rapport du destinataire contenant des statistiques comme le nombre de paquets reçus, le taux de perte, la gigue. RTCP qui est facultatif met à disposition des donnés supplémentaires permettant d'affiner le traitement de la QoS d'une session RTP. En exploitant les informations RTCP, les nœuds réseau peuvent calculer les facteur R et l'E-model définis dans la recommandation G.107 afin d'affecter une note MOS à la communication concernée.

RTP pour transporter le media

RTP
- identifie le type d'information transportée **(ex : le type de codeur voix dans le champ PT)**
- ajoute des marqueurs temporels aux paquets
- ajoute des numéros de séquence,
- permet de réordonner les paquets, ou tenir compte de leur arrivée dans le désordre (ex : chaque paquet vidéo pourra être décodé et placé au bon endroit sur l'écran, sans attendre ses prédécesseurs)
- le récepteur est informé de la perte de paquets (Il peut alors mettre en œuvre des mécanismes permettant de compenser ou réparer la perte d'information : redondance, approximation, …)
- le récepteur est informé de la gigue (Il peut alors mettre en œuvre des mécanismes afin de compenser la gigue : modification taille buffer gigue dynamique)

Type de Payload (PT) i.e le codec qui est effectivement utilisé

Quelques types statiques de Payload sont assignés par l'IANA



Payloads statiques

PT	Nom codec	Audio/Video	Fréq. Hz	RFC
0	PCMU	A	8000	[RFC3551]
1	Reserved			
2	Reserved			
3	GSM	A	8000	[RFC3551]
4	G723	A	8000	[Vineet_Kumar][RFC3551]
5	DVI4	A	8000	[RFC3551]
6	DVI4	A	16000	[RFC3551]
7	LPC	A	8000	[RFC3551]
8	PCMA	A	8000	[RFC3551]
9	G722	A	8000	[RFC3551]
10	L16	A	44100	[RFC3551]
11	L16	A	44100	[RFC3551]
12	QCELP	A	8000	[RFC3551]
13	CN	A	8000	[RFC3389]
14	MPA	A	90000	[RFC3551][RFC2250]
15	G728	A	8000	[RFC3551]
16	DVI4	A	11025	[Joseph_Di_Pol]
17	DVI4	A	22050	[Joseph_Di_Pol]
18	G729	A	8000	[RFC3551]
19	Reserved	A		
25	CelB	V	90000	[RFC2029]
26	JPEG	V	90000	[RFC2435]
27	Unassigned	V		
28	nv	V	90000	[RFC3551]
31	H261	V	90000	[RFC4587]
32	MPV	V	90000	[RFC2250]
33	MP2T	AV	90000	[RFC2250]
34	H263	V	90000	[Chunrong_Zhu]
35-71	Unassigned	?		
72-76	Reserved for RTCP conflict avoidance			[RFC3551]
77-95	Unassigned	?		
96-127	dynamic	?		[RFC3551]

https://www.iana.org/assignments/rtp-parameters/rtp-parameters.xhtml

Dégradation du media (et de la signalisation dans un moindre mesure)

Latence (engl. Lag) : temps mis par un paquet émis par une origine pour atteindre sa destination, délais de traitement et d'acheminement compris. Correspond au demi RTD couramment utilisé en IP. Le recommandation G.114 stipule que le délai de transmission de « bouche à oreille » (concerne donc le media) doit être inférieur à 150ms.
Dans les faits, un délai supérieur à 200ms commence à être perceptible par les locuteurs et affecte l'interactivité des conversations.

Qualité perçue et délai :

- 0 à 150 ms : communications fluides
- 150 à 300 ms : communications moyennement à faiblement interactives
- 300 à 700 ms : correspond à des communications half duplex
- > 700 ms : communication impossible

Gigue (engl. Jitter) : variation de la latence. C'est un phénomène courant lié à la transmission asynchrone et au multiplexage statistique (file d'attente). Les effets de la gigue peuvent être atténués en plaçant une mémoire tampon du côté du récepteur. Ce traitement, permettant de stabiliser la variation de délai entre 2 paquets consécutifs et remettre en ordre les paquets éventuellement dé-séquencés dans le réseau. Ceci augmente la latence globale et potentiellement le délais inter paquet.

Dégradation du media (et de la signalisation dans un moindre mesure)

Perte de paquets (engl. Drop) : disparition d'information généralement liée à :

- une congestion temporaire dans les réseaux,
- une défaillance d'un équipement participant à la chaine de transmission (yc terminaux),
- une mauvaise gestion de la priorisation des paquets RTP dans les files d'attente.

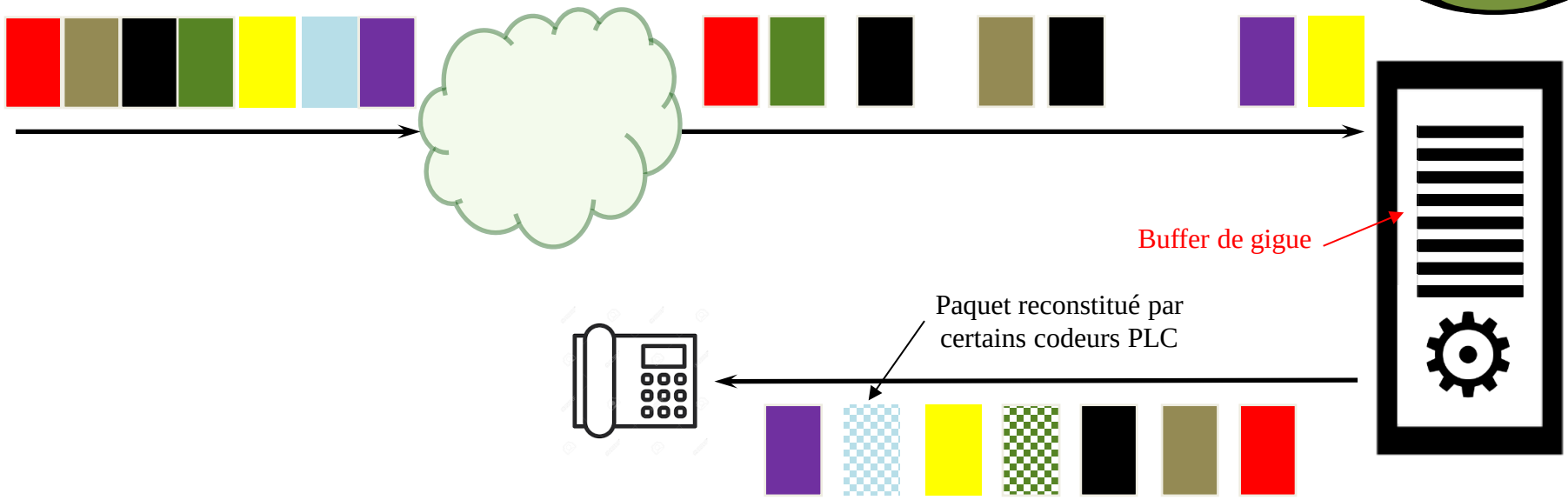
Les codeurs permettent de reconstituer les données manquantes en se basant sur les contenus précédemment reçus (PLC).

Bruit (engl. Noise) : les imperfections introduites lors de la numérisation du signal analogique origine ne pourront pas être éliminées. Par ailleurs, le transcodage numérique éventuel du media dans certains nœuds réseau intermédiaires accroit la latence et est susceptible d'altérer le media.

Latence, gigue et perte de paquets sont les 3 principales raisons qui contribuent à la dégradation de la qualité vocale. Des mécanismes existent pour en atténuer les effets. Le bruit introduit lors de la conversion numérique du signal analogique origine ainsi que les altérations liées aux transcodages éventuels ne pourront pas être corrigés.

Lorsque la VAD est activée le phénomène appelé « clipping » peut survenir. Il s'agit d'une anomalie de détection du niveau d'activité vocale lors des transitions parole <-> silence. Ce phénomène particulièrement audible est très désagréable.

Principe du buffer de gigue (jitter buffer)



Le rôle du buffer de gigue est de présenter un flux de paquets relativement stable à la fonction de traitement du media. Cette fonction peut être mise en œuvre de façon optionnelle en réception de trafic RTP et permet dans un intervalle de temps donnée de :

1. **Supprimer les paquets dupliqués,**
2. **Remettre en ordre les paquets reçus hors séquence,**
3. **Rendre quasiment constant le délai entre 2 paquets consécutifs émis,**
4. **Les paquets arrivant en dehors de la fenêtre de temps « droppés »,**

Comme le traitement consiste à retenir pendant une courte durée Δt tous les paquets reçus, la latence va se trouver augmentée de Δt. Il existe 2 implémentations de buffer de gigue statique (Δt fixe) ou dynamique qui s'adapte automatiquement au délai réseau mesuré par la machine (Δt variable). Mécanisme de buffer de gigue à mettre en œuvre dans les terminaux plutôt que dans les réseaux.

Fonctionnement buffer de gigue (jitter buffer)

Initialisation: au démarrage, le buffer de gigue se remplit à concurrence de ’n’ paquets RTP (‘x’ est la variable d’initialisation du buffer exprimée en msec ou octet permettant de déterminer la valeur de ‘n’).

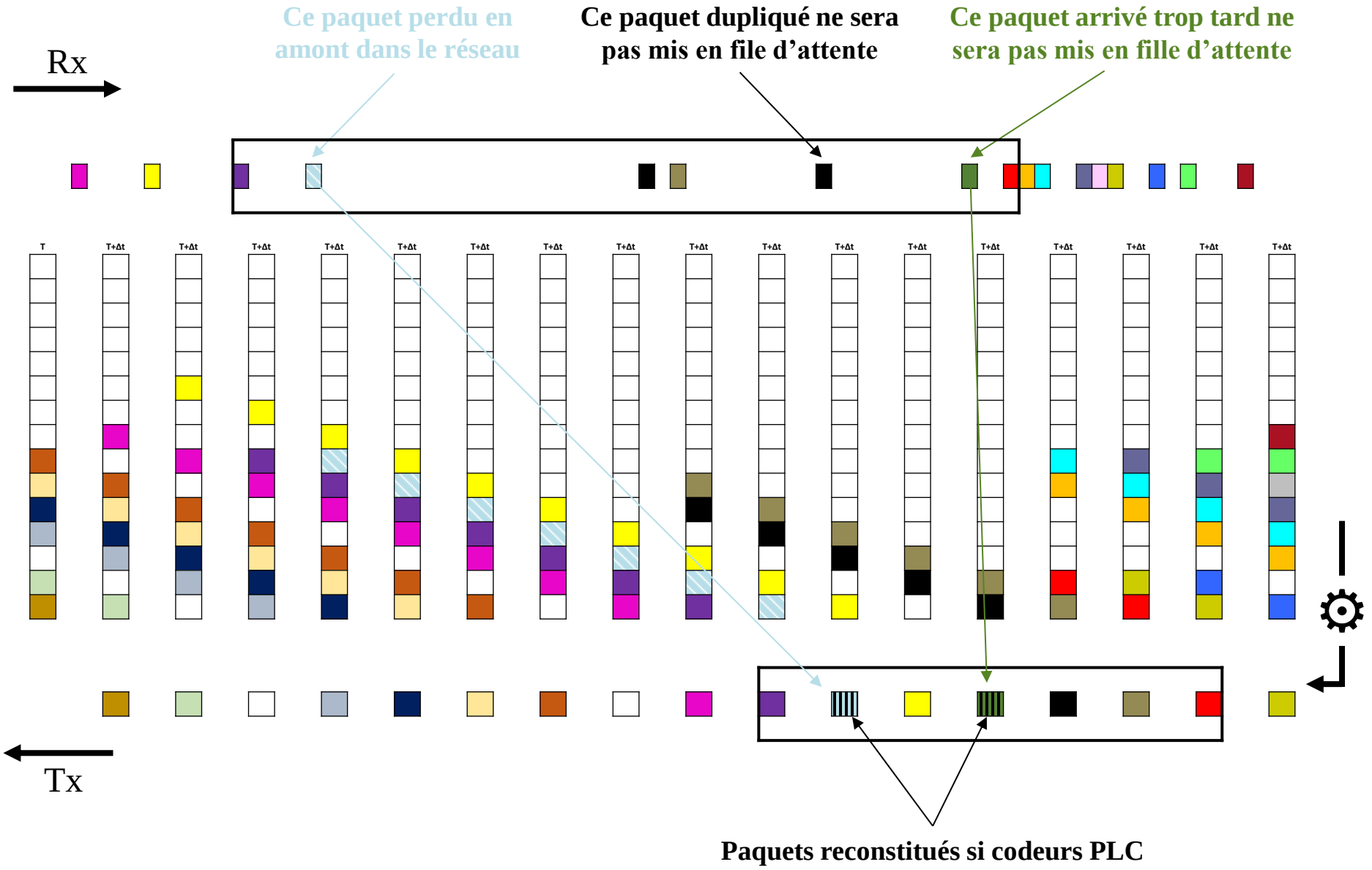
Fonctionnement: la réception du paquet RTP dont le numéro de séquence est >= ‘n’, déclenche la remise à l’applicatif du premier paquet se trouvant dans la file d’attente. La file d’attente se vide ensuite à une cadence Δt.

Tout paquet hors séquence arrivant dans la fenêtre (i.e. dont le numéro de séquence est > au plus petit numéro de séquence présent dans la file et < au plus petit numéro de séquence présent dans la file + n) est placé dans le buffer de gigue, sinon il est droppé.

Tout paquet perdu en amont l’est définitivement, des mécanismes prédictifs implémentés au niveau de certains codeurs (PLC) peuvent tenter de générer un paquet de remplacement.

Conséquence: l’implémentation du buffer de gigue au niveau d’un équipement augmente le délai global de transmission de Δt (une fois le buffer stabilisé après la période de convergence liée à la phase d’initialisation).

Traitement


Ce paquet perdu en amont dans le réseau
Ce paquet dupliqué ne sera pas mis en file d'attente
Ce paquet arrivé trop tard ne sera pas mis en fille d'attente
Rx
T
T+Δt
Tx
Paquets reconstitués si codeurs PLC

Phénomène d'écho

L'écho est défini comme un signal (quasi identique à celui reçu) renvoyé avec décalage et atténuation vers l'émetteur. Il se superpose au signal utile. L'oreille humaine n'est pas perturbée lorsque le flux vocal d'un utilisateur lui revient en moins de 30 ms. Au-delà, le phénomène d'écho devient audible puis vraiment gênant si le temps de boucle est supérieur à 50 ms auquel cas l'emploi d'un dispositif de suppression d'écho s'impose.

Il existe deux types d'écho :

- Electrique : il est généré par l'interface 2/4 fils qu'on trouve dans les commutateurs d'abonné ou PABX. La désadaptation d'impédance liée à la conversion 2/4 fils agit comme un « obstacle » qui renvoie une partie du signal vers celui qui l'a émis.

- Acoustique : se produit lorsque le terminal du distant renvoie le flux audio diffusé par son propre haut-parleur lorsque la fonction mains libres est activée.

Les annuleurs d'écho étaient systématiquement insérés dans les réseaux longue distance TDM puisque ceux-ci introduisent des délais de transit quelquefois importants. L'utilisation du transport IP a aggravé la situation puisque les délais de transit introduit par le transport de la voix sur réseaux IP sont toujours plus importants qu'en TDM.

L'écho électrique est occasionné par le réseau tandis que l'écho acoustique est causé par les terminaux.

Techniques d'annulation de l'écho

Pour protéger l'utilisateur A il faut implémenter un mécanisme d'annulation d'écho au plus proche de sa source, c'est-à-dire à proximité de B, et vice versa.

Idéalement les fonctionnalités d'annulation d'écho doivent être présentes dans les terminaux. Cependant, ce n'est pas toujours suffisant notamment lorsque l'écho se produit au niveau du réseau ou d'un point d'accès au réseau.

L'écho est caractérisé par son délai par rapport au signal d'origine et par le niveau du signal réfléchi. Gêne = f(délai,niveau)

L'écho électrique est en voie de disparition puisque les terminaux analogiques se raréfient avec la disparition du RTC. L'écho généré par la conversion numérique / analogiques est supprimé par les PBX ou les box internet, par contre il n'est pas traité par les commutateurs d'abonnés RTC. Pour annuler l'écho électrique, outre une adaptation d'impédance appropriée on utilise une technique numérique qui compare les signaux reçu et émis dans une fenêtre de traitement de taille déterminée (32, 64 voire 128 ms). Le signal émis similaire au signal reçu sera éliminé. Les réseaux RTC « legacy » doivent maintenant gérer l'annulation d'écho à leur niveau car les réseaux longue distance on presque tous migré vers le transport IP.

L'écho acoustique est bien plus compliqué à annuler que l'écho électrique car le signal renvoyé est perturbé par le bruit de fond. Un microphone plus directif implémenté dans les terminaux, la baisse du volume des haut-parleurs et/ou l'utilisation d'écouteurs améliorent sensiblement les choses. L'emploi d'algorithmes de filtrage adaptatif directement implémentés dans les terminaux, PBX et box modernes permet d'annuler l'écho acoustique plus efficacement qu'auparavant.

Principe de l'annuleur d'écho

Comme indiqué plus haut, c'est le dispositif de gestion de l'écho installé côté distant qui améliore le confort auditif de l'abonné local. La fenêtre d'analyse des annuleurs étant relativement étroite, ceux-ci doivent être installés au plus proche de ce qui provoque la réflexion du signal incident.

L'annuleur d'écho AE1 effectue les opérations suivantes :

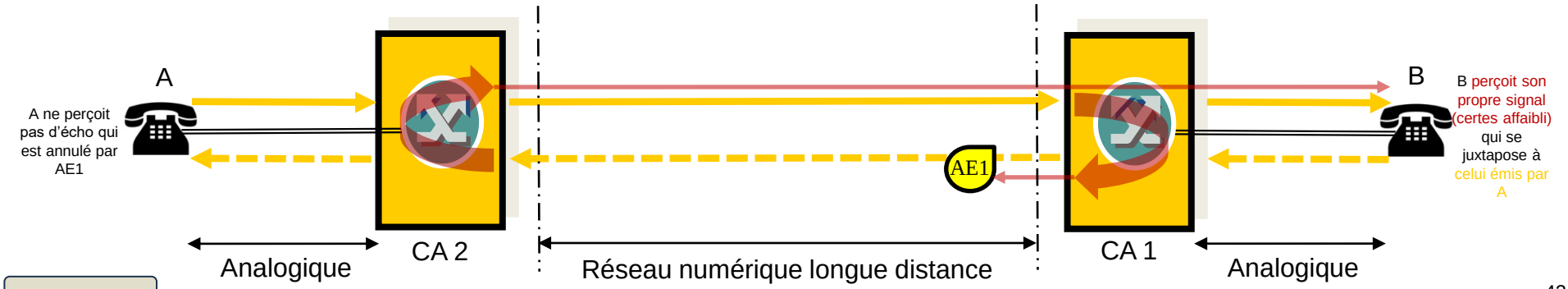
1. mémorisation du signal incident reçu de A pendant un laps de temps appelé temps de convergence,
2. estimation de l'écho par comparaison du signal reçu de B à émettre vers A et le signal mémorisé,
3. soustraction de l'écho estimé au signal présent sur la voie d'émission vers A,

S'il existe à proximité de CA 2, l'annuleur d'écho AE2 effectue les opérations suivantes :

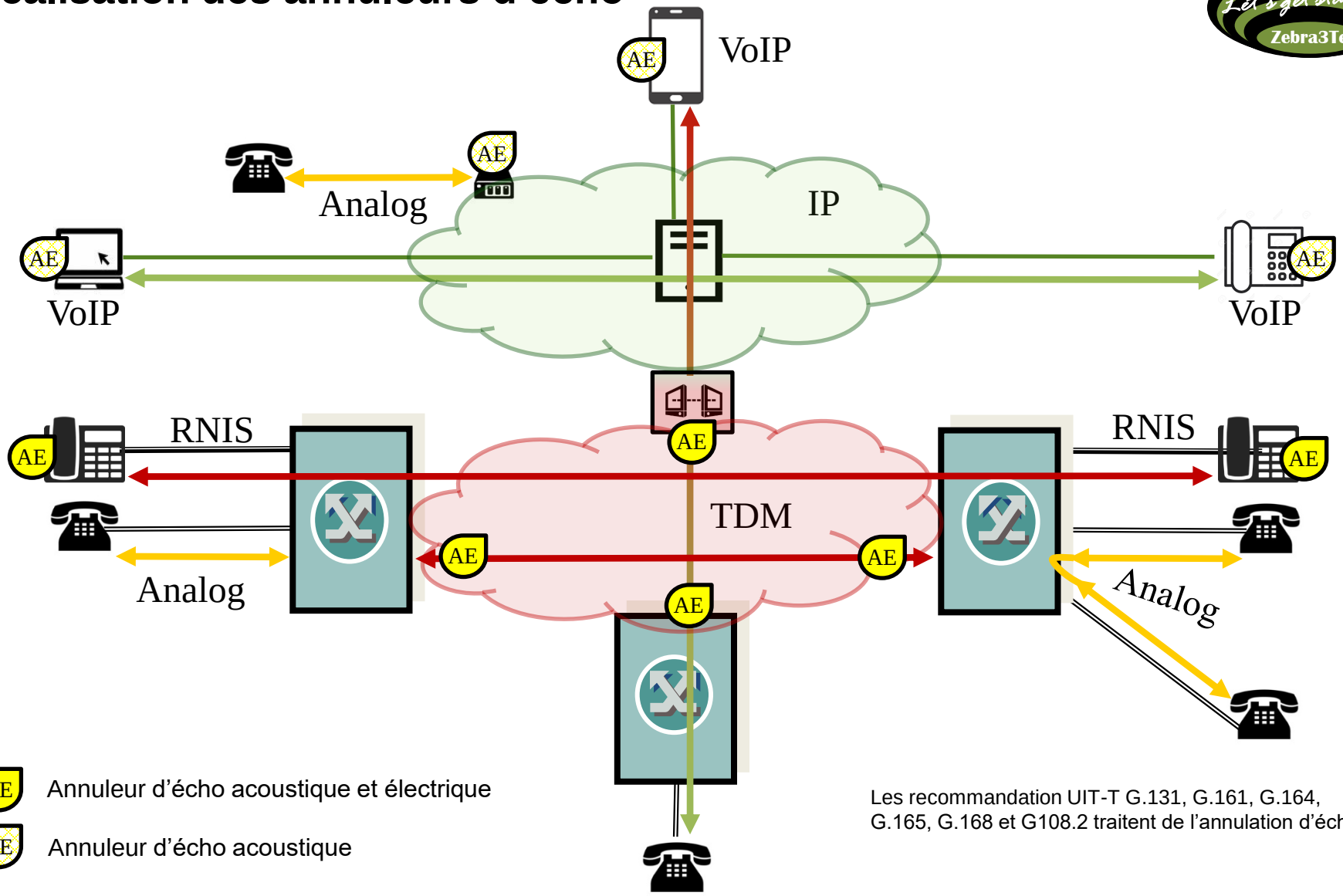
1. mémorisation du signal incident reçu de B pendant un laps de temps appelé temps de convergence,
2. estimation de l'écho par comparaison du signal reçu de A à émettre vers B et le signal mémorisé,
3. soustraction de l'écho estimé au signal présent sur la voie d'émission vers B,

N.B.: tout écho éventuellement contenu dans le signal reçu de B à émettre vers A ne sera pas annulé par AE1 s'il est reçu hors fenêtre d'analyse (idem pour AE2).

CA 1 et CA 2 génèrent de l'écho électrique lié à l'adaptation 4/2 fils. AE1, implémenté un peu en aval de CA 1 protège A de l'écho généré par CA 1. B n'est pas protégé.



Localisation des annuleurs d'écho



AE Annuleur d'écho acoustique et électrique

AE Annuleur d'écho acoustique

Les recommandation UIT-T G.131, G.161, G.164, G.165, G.168 et G108.2 traitent de l'annulation d'écho

Commutateur TDM

GateKeeper H.323 ou Proxy SIP

Passerelle IP/TDM

Adapteur de Terminal Analogique ou box FAI ou PBX

Mesure de la QoS

On a vu précédemment que la QoS était affectée par les phénomènes suivants :

1. Latence (engl. Lag) dont les facteurs contributifs sont :
   - Le délai d'échantillonnage
   - Le délai de transport
   - Le délai des buffers de gigue
2. Gigue (engl. Jitter)
3. Perte de paquets (engl. Drop)
4. Bruit (engl. Noise)

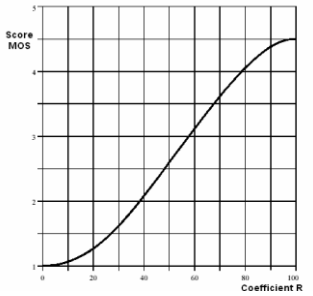
Indicateurs de la QoS

La MOS (ITU P800) est une mesure subjective (1 < MOS < 5) de la QoS voix basée l'évaluation de la qualité de la conversation auprès d'opérateurs humains.

L'ETSI a développé l'E-model (ETR 250) afin de déterminer la qualité du transport de la voix de bout en bout (normalisé ensuite par l'ITU G.107). L'objet de ce modèle est de calculer une grandeur appelée facteur R en fonction des sources de dégradation énoncées précédemment. (0 < R < 100)

R = R0 – Is – Id – Ie + A

R0 : capital initial de QoS = 94,3 en VoIP.
Is : dommages simultanés avec l'émission de la voix (bruit de fond)
Id : dommages dus au délai de transmission et de transport (gigue, perte paquet, …)
Ie : dommage de distorsion causés par les équipements (Codec, transcodage, buffer, ...)
A : coefficient d'amélioration (réseau fixe ou mobile)



La conversion MOS = f(R) est fournie dans le graphe ci-dessus. On considère que la qualité est bonne ou très bonne lorsque le Facteur R est > 80 c'est dire quand la note MOS est supérieure à 4.

Le tableau ci-contre présente l'influence des principaux facteurs réseaux sur la qualité vocale d'une communication VoIP en codec G.711 64 kbps

QoS	BONNE si	MAUVAISE si
Latence	L < 150 ms	L > 400 ms
Gigue	et G < 20 ms	ou G > 50 ms
Perte Paquet	et PP < 1%	ou PP > 3%

DTMF

Les tonalités DTMF (ou fréquences vocales) sont générées par des terminaux quand on presse une touche pendant un minimum de 50ms. Elles sont la combinaison de 2 fréquences pures haute et basse émises simultanément qu'il est peu probable de retrouver dans la parole humaine. Ces DTMF sont fréquemment utilisées pour interagir avec des serveurs vocaux.

(Hz)	1 209	1 336	1 477	1 633
697	1	2	3	A
770	4	5	6	B
852	7	8	9	C
941	*	0	#	D

Les codeurs bas débit (autres que G.711) ne sont pas appropriés pour véhiculer ces tonalités DTMF car les codages/décodages successifs les déformeront à tel point qu'elles ne seront pas reconnues par le distant.

DTMF

Possibilités de transport des DTMF :

- dans la bande : si codeur G.711, les DTMF peuvent être véhiculés de façon banalisée dans des paquets RTP au même titre que la voix,

- RFC4733 (RFC2833 auparavant) : l'information DTMF est envoyée dans le flux RTP sous forme d'un événement et non pas d'un signal codé comme précédemment. La conservation des séquencement et référence temporelle permet de garder la synchronisation entre les flux media et DTMF. La RFC4733 spécifie un format de paquet RTP pour représenter les informations DTMF. L'entête du paquet RTP contiendra l'information (payload type) indiquant que le paquet RTP transporte une tonalité DTMF,

- en SIP, les méthodes INFO et SUSCRIBE/NOTIFY permettent d'envoyer des événements téléphoniques hors bande (i.e. au niveau signalisation). L'interopérabilité problématique des différentes implémentations SIP rend peu fiable l'utilisation de ce mécanisme si les éléments réseaux traversés ne proviennent pas du même fournisseur,

Il est fortement conseillé d'utiliser le mécanisme RFC2833/RFC4733 pour le transport des DTMF car elle garantit la synchro media/DTMF d'une part et l'interopérabilité entre systèmes hétérogènes d'autre part

Early-Media

Il s'agit d'un mécanisme permettant à 2 UA SIP d'échanger du media avant que la communication soit effectivement établie.

- L'early media se réfère à un flux media envoyé après l'émission d'un message INVITE qui n'a pas encore reçu de réponse finale,

- Il peut être mono-directionnel ou bi-directionnel,

- L'early media peut principalement se gérer comme suit :
  - De façon implicite, un UA se met à l'écoute du média sur le port média indiqué dans le SDP qu'il émet. Par conséquent, le serveur qui reçoit ce SDP peut potentiellement immédiatement envoyer du média à destination ce cet UA,
  - De façon explicite, en valorisant à « supported » l'entête P-Early-Media (RFC 5009) dans les messages INVITE qu'il émet et en paramétrant l'entête P-Early-Media des réponses 18x (avec SDP) émises par le serveur à la valeur « sendonly » ou « sendrecv ».

Les arguments TDM / VoIP

	TDM	VoIP
COUTS		
Equipement		
Transport		
RH Expertise		
OPERATION		
Mise au point		
Convergence voix/données		
Mise en œuvre réseau		
Coordination équipes		
FONCTIONNEL		
Flexibilité		
Virtualisation		
Intégration des services		

	TDM	VoIP
TRANSPORT		
Délais		
Gigue		
Perte d'information		
MEDIA		
Codage voix haute définition		
Sensibilité au transport		
Qualité vocale		
SIGNALISATION		
Implémentation des standards		
Interfonctionnement		
Evolutivité des standards		
Richesse des standards		
SECURITE		
Fraude		
Vulnérabilité aux attaques		
Déni de service		

L'introduction de la VoIP a permis de sensiblement réduire le coût des infrastructures réseau pour un service équivalent / TDM. La facilité de développer des nouvelles fonctionnalités est le principal atout de la VoIP malgré un certain nombre de facteurs techniques défavorables

3. Le protocole SIP (RFC3261)

SIP généralités

SIP (Session Initiation Protocol) a été défini pour la première fois en 1999 dans la RFC 2543 par le groupe de travail MMUSIC afin de donner un cadre technique aux communications multimédia. Basé sur les protocoles SDP et SIP, la RFC 2543 utilise le protocole RTP (RFC 1889) pour le transport du media.

Le contexte du début des années 2000 à fortement favorisé SIP au détriment de H.323 pourtant à peine plus vieux. A cette époque de nombreuses start-up ont massivement investi pour promouvoir SIP même si l'interopérabilité avec le RTC s'est avérée bien plus problématique que prévu.

L'objectif de SIP est de rester le plus simple possible.

En 2002 la nouvelle **RFC 3261** a « compilé » les multiples améliorations apportées à la RFC précédente qui comptait de nombreuses lacunes. Avec cette nouvelle version, SIP s'est considérablement complexifié en pointant vers d'autres documents de référence comme les RFC 3262, 3263, 3264, 3265 et 3266. SIP utilise des concepts et éléments des protocoles internet HTTP (design client-serveur, URL/URI) et SMTP (entêtes, format d'encodage texte).

Il est à noter que la première version 2.0 du protocole SIP définie par la RFC 3261 dès son origine est toujours en vigueur.

Transport du protocole SIP

Au sens du modèle OSI (Open System Interconnexion), SIP est une application orientée session de niveau 5 qui peut être transportée sur une couche de niveau 4 UDP, TCP, TLS ou SCTP. En général SIP est utilisé sur transport UDP.

Selon la RFC 3261et l'IANA les ports SIP par défaut sont :

- 5060 pour SIP
- 5061 pour SIPS (SIP / TLS)

L'utilisation des ports SIP ou SIPS par défauts n'est pas obligatoire, d'autres ports (non réservés) peuvent être choisis pour établir des sessions SIP ou SIPS.

Quant au flux média, il peut être transporté sur UDP ou TCP. On peut aussi utiliser le protocole SRTP si on souhaite le chiffrer. Il est fortement déconseillé de véhiculer le flux media sur TCP, option qui est d'ailleurs rarement disponible dans les UA SIP.

	TCP	UDP
POUR	Support des MTU > 1500	Etablissement plus rapide des communications
	Fréquence moindre des keep alive pour maintenir les ports NAT ouverts	Moins d'overhead
CONTRE	Plus de puissance CPU nécessaire notamment quand le trerminal s'enregistre fréquemment	Les problème de fragmentation sont un facteur limitant dans les cas d'appel complexes
	Latence aditionnelle dans le temps d'établissement d'un appel	Fréquence accrue des keep alive peut décharger plus rapidement la batterie des terminaux mobiles

Eléments importants (*obligatoires) des messages SIP

Un message SIP est structuré de la façon suivante :

- **Request ou Status Line :**
  indique le type de message utilisé et comprend l'adresse du destinataire (sous forme d'URI) à qui le message est destiné (pour les requêtes), SIP/2.0 est la version SIP.

  Eléments de langage :
  - Request-Line : INVITE sip:2222@10.20.30.40:5070;user=phone SIP/2.0
  - Request-URI : INVITE sip:2222@10.20.30.40:5070;user=phone SIP/2.0
  - Request-URI user : INVITE sip:2222@10.20.30.40:5070;user=phone SIP/2.0
  - Request-URI host : INVITE sip:2222@10.20.30.40:5070;user=phone SIP/2.0

- **Message Header :**
  c'est la partie du message SIP qui contient toutes les informations utiles à l'établissement d'une session SIP. Ces informations sont structurées sous forme d'entêtes normalisées par des RFC,

- **Message Body :**
  cette partie, optionnelle, permet de véhiculer des informations complémentaires. Le type de contenu et la longueur du « Message Body » sont indiqués dans les entête « Content-Type » et « Content-Length » la section « Message Header ». Il existe d'autres entêtes Content-xxxxx pour apporter des précisions sur la disposition, l'encodage, le langage du contenu des sous-sections. A titre d'exemple, l'annonces des fonctionnalités d'un UA, la négociation des capacités media (SDP), l'encapsulation d'éléments protocolaires (ISUP), … sont des informations qu'on retrouve dans la partie « Message Body » des certains message SIP.

Entêtes importants (*obligatoires) des messages SIP

***Via** contient l'adresse de l'entité et le port SIP vers lesquels les réponses devront être envoyées. Un champ via est inséré à chaque fois qu'un élément réseau SIP est traversé. Un élément SIP route une réponse en exploitant le contenu du « Via » et retire son propre entête « Via » avant d'envoyer la réponse. La transaction SIP est identifiée par le paramètre « branch »,

***Max-Forward** est inséré par l'UAC et est décrémenté d'une unité à chaque passage dans un nœud SIP. Si la valeur 0 est atteinte, un message d'erreur sera généré. (Fonctionnement similaire au TTL pour prévenir les bouclages),

***From** identifie l'appelant, il doit contenir un paramètre tag qui est un identificateur aléatoire. Les réponses retournées par l'appelé doivent inclure un paramètre tag dans l'entête « To »,

***To** identifie la cible, les réponses retournées par l'UA appelé doivent inclure un paramètre « tag » aléatoire dans l'entête « To » et mettre dans l'entête « From » la valeur du « tag » précédemment reçu dans le « From »,

***Call-ID** contient l'identificateur unique généré par l'initiateur de l'appel valable pour toute la durée de la communication. Un dialogue SIP est identifié sans ambigüité grâce à la valeur du Call-ID,

***Cseq** permet de corréler les requêtes aux réponses correspondantes. C'est un numéro de séquence aléatoire suivi du nom de la méthode SIP à laquelle il s'applique. Le numéro est incrémenté à chaque nouvelle méthode SIP émise sauf requêtes « ACK » et « CANCEL ». Différencie les retransmissions des nouveaux messages,

Content-Length si différent de 0, indique qu'un body pouvant par exemple contenir un SDP (voire un message ISUP dans le cas du SIP I) est à suivre,

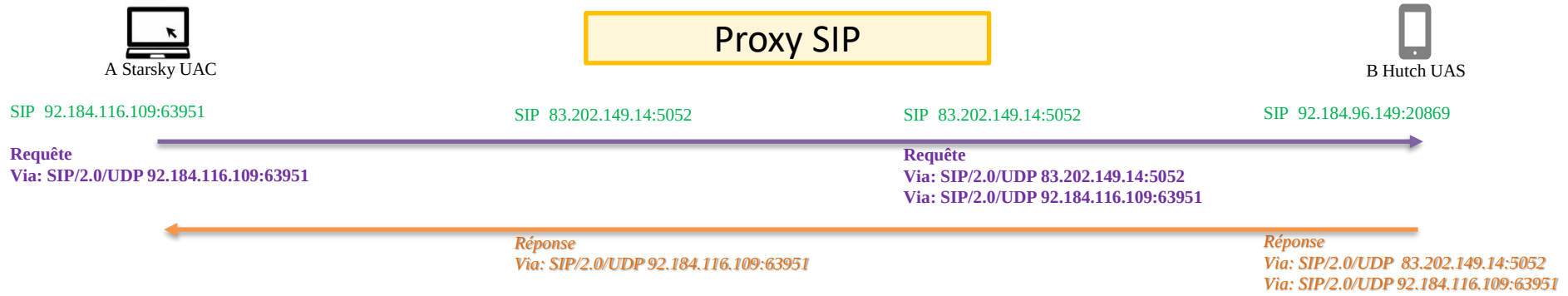
Content-Type annonce le type de contenu présent dans la section « Message Body », on retrouve souvent « Content-Type: application/sdp » pour la négociation des capacités media entre UA SIP,

Session-Expires durée de la sessions SIP négociée lors de l'établissement de la communication SIP, un re-INVITE est émis avant la fin de cette durée afin de « rafraichir » la session SIP en cours (à ne pas confondre avec l'entête Expires qui caractérise l'enregistrement),

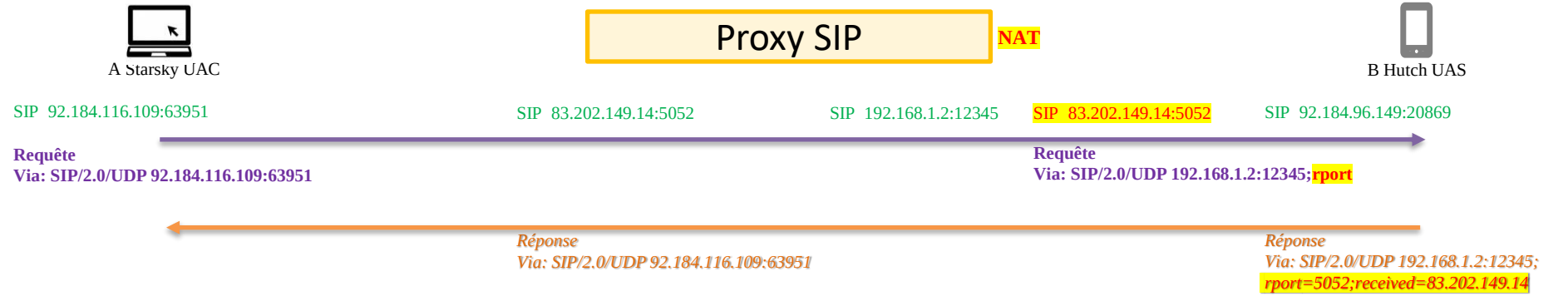
Expires durée d'enregistrement d'un terminal ou trunk SIP. A la moitié de cette durée, un nouveau message REGISTER est émis par l'UA qui s'est enregistré pour vérifier que l'UA lui faisant face est fonctionnel.

Utilité du header Via

Le rôle du header via est de définir le chemin que l'UA devra utiliser pour router ses réponses. Chaque instance SIP traversée est théoriquement censée insérer dans les requêtes SIP qu'il émet son propre header via "au dessus" de ceux qui existent déjà. Inversement, dans les réponses SIP, chaque instance SIP traversée supprime son propre header via.



Lorsqu'un device SIP se trouve derrière un NAT il y est probable qu'il insère une IP non routable dans le header via qu'il va générer. La RFC 3581 propose un solution pour y remédier à l'aide des paramètres "rport" et "received" du header via.



Si le paramètre "rport" est présent dans le header via reçu par B, pour router ses réponses B utilisera les IP et port contenus dans l'entête IP de la requête reçue (plutôt que dans le header via). Il peut les mentionner respectivement dans les paramètres "received" et "rport" du header via de la réponse qu'il transmet vers l'émetteur A de la requête. Les IP et port destination de la réponse émise par B vers A correspondent aux IP et port source contenus dans l'entête IP de la requête reçue par B.

Syntaxe des messages SIP

Les messages SIP sont encodés en utilisant une syntaxe similaire à HTTP/1.1 (RFC2068 et 2616) en se basant sur un modèle transactionnel client/serveur. Le jeu de caractères utilise le codage UTF-8 (RFC2279), les lignes se terminent par CR/LF.

Requête / Méthode	RFC	Description
INVITE	3261	Etablissement d'une session de communication entre UA
ACK	3261	Confirmation de la réception d'une réponse finale suite à l'émission d'un INVITE
OPTIONS	3261	Récupération des informations relatives aux capacités d'un UA distant, sans pour autant établir d'appel
BYE	3261	Libération d'une session SIP établie
CANCEL	3261	Annulation le process d'établissement d'une communication (session SIP non encore été établie)
REGISTER	3261	Enregistrement, ré-enregistrement ou désenregistrement (*Expires*: *0)* d'un UA auprès d'un registrar
INFO	2976	Echange de signalisation comme les tonalités DTMF pendant la session
MESSAGE	3428	Emission de messages instantanés
NOTIFY	3265	Envoi de notifications d'événements
PRACK	3262	Fiabilisation de la transmission des réponses provisoires en acquittant leur réception
REFER	3515	Redirection d'un appel vers un autre UA
SUSCRIBE	3265	Abonnement à la réception de notifications lorsqu'un événement souscrit se produit
UPDATE	3311	Mise à jour des paramètres d'une session multimédia

2 types de messages sont utilisés en SIP :

1. **les requêtes ou méthodes,**
2. **les réponses,**

Une nouvelle requête INVITE alors que la session SIP est établie s'appelle un re-INVITE.
Le re-INVITE indique généralement que certains descripteurs de l'INVITE initial ont été modifiés.

Une session SIP est établie lorsque l'émetteur de l'INVITE acquitte (avec la méthode ACK) la réponse finale reçue.

Syntaxe des messages SIP

Un UA SIP peut répondre à une requête SIP à l'aide de messages réponses NXX avec 1<= N <=6, N définissant la catégorie de réponse.
Toutes les réponses sont finales exceptées celles qui commencent par 1 appelées réponses provisoires, c'est-à-dire qu'elles ne terminent pas la transaction courante.
La première ligne d'une réponse débute par un code numérique suivi d'une brève explication. En fonction de la réponse, des données supplémentaires comme une description de session SDP peuvent être présentes.
Les réponses peuvent contenir un entête optionnel « Reason » (RFC3326). Celui-ci permet de véhiculer de façon transparente les indications de cause Q.850 dans le cas d'interfonctionnement avec les réseaux TDM par exemple.

Une extension au protocole SIP pour acquitter les réponses provisoires 1xx utilise l'étiquette d'option 100rel et définit la méthode d'accusé de réception PRACK (*Provisional Response ACKnowledgement*).

Réponses	RFC	Description
1xx	3261	Information : requête reçue et en cours de traitement, la transmission des réponses 1xx n'est pas fiable
2xx	3261	Succès : requête reçue, comprise et acceptée
3xx	3261	Redirection : autre action requise pour compléter le traitement de la requête
4xx	3261	Erreur du client : requête mal formatée ou ne pouvant pas être exécutée par le serveur
5xx	3261	Erreur serveur : échec du traitement d'une requête apparemment valide, problème de syntaxe
6xx	3261	Echec global : requête invalide ne pouvant être traitée par aucun serveur

Temporisateurs (timers) SIP

Temporisateur	Valeur par défaut	Signification
T1	500 ms	Retransmission INVITE (la valeur de la tempo double à chaque tentative - max 8 tentatives = 32s -)
T2	4 sec.	Intervalle de retransmission maximale pour les requêtes non-INVITE et les réponses INVITE
T4	5 sec.	Durée maximale pendant laquelle un message peut rester dans le réseau
Temporisateur A	initialement T1	Intervalle de retransmission des requêtes INVITE, pour UDP uniquement
Temporisateur B	64*T1	Temporisateur du délai pour opérations INVITE
Temporisateur C	> 3 min.	Délai d'expiration des opérations Proxy INVITE
Temporisateur D	> 32 sec. pour UDP	Délai d'attente pour les retransmissions de réponse
Temporisateur E	initialement T1	Intervalle de retransmission des requêtes non-INVITE, UDP uniquement
Temporisateur F	64*T1	Temporisateur du délai pour opérations non-INVITE
Temporisateur G	initialement T1	Intervalle de retransmission des réponses INVITE
Temporisateur H	64*T1	Délai d'attente pour le reçu d'accusé de réception (ACK)
Temporisateur I	T4 pour UDP	Délai d'attente pour les retransmissions d'accusés de réception (ACK)
Temporisateur J	64*T1 pour UDP	Délai d'attente pour les retransmissions des requêtes non-INVITE
Temporisateur K	T4 pour UDP	Délai d'attente pour les retransmissions de réponse

Mécanisme de retransmission des INVITE ne recevant aucune réponse :

A émet un premier message INVITE vers B à l'instant T0.

Si A n'a pas reçu de réponse à l'instant $T0+T1x2^{(n-1)}$, A réémet le même message INVITE n fois en doublant la valeur du temporisateur T1 précédent tant que $T0+\Delta t <=$ Temporisateur B

Tentative n	0	1	2	3	4	5	6	7	8
Tempo = $T1 \times 2^{(n-1)}$	0	$T1 \times 2^{(1-1)}$	$T1 \times 2^{(2-1)}$	$T1 \times 2^{(3-1)}$	$T1 \times 2^{(4-1)}$	$T1 \times 2^{(5-1)}$	$T1 \times 2^{(6-1)}$	$T1 \times 2^{(7-1)}$	$T1 \times 2^{(8-1)}$
Δt si T1 = 0,5 s	0	0,5	1	2	4	8	16	32	64

Par exemple : T1=0,5s alors Temporisateur B=32s. La 8^e tentative (ou 7^e re-tentative) sera la dernière. Pour éviter les retransmissions intempestives, le bon usage veut que les récepteurs d'INVITE envoient immédiatement la réponse 100 Trying. Cette réponse provisoire signifie que l'INVITE a bien été reçue et est en cours de traitement. La réception de ce message permet à A de désarmer T1 et attendre les messages subséquents.

Enregistrement des UA SIP

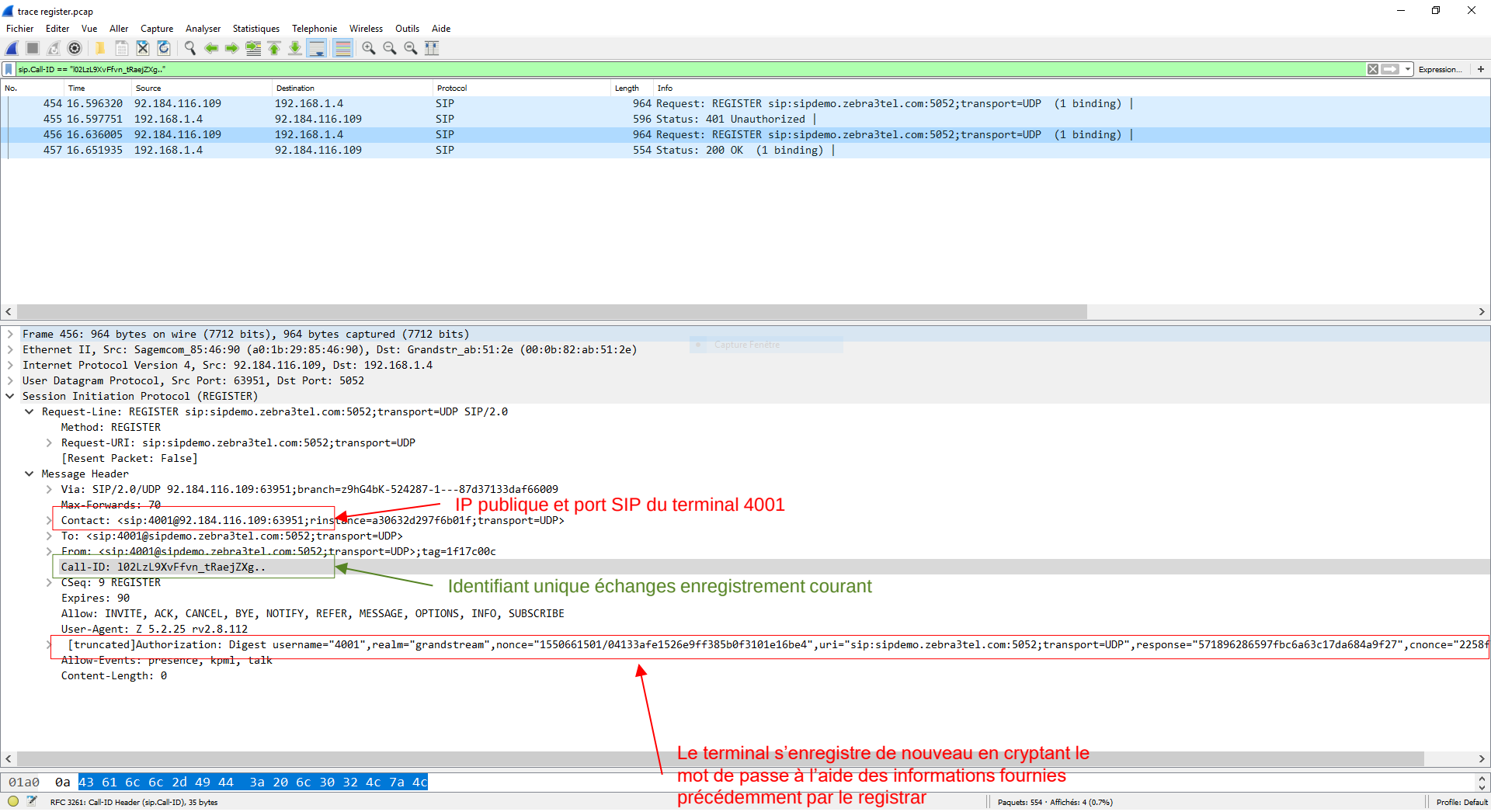
Un terminal SIP (ou User Agent) utilise la requête «REGISTER» pour indiquer à la base de données centralisée des abonnées (Registrar) quels sont ses adresse IP et port SIP. La phase d'enregistrement est primordiale car elle permet aux fournisseurs de services SIP d'authentifier l'UA SIP souhaitant s'enroler. Une fois authentifié, l'UA sera autorisé à émettre et recevoir des appels. Pour se dé-enregistrer il faut donner la valeur 0 au header Expires.

Principales informations se trouvant dans les entêtes du REGISTER :

- Le header Via transporte l'IP de l'initiateur de la requête puis celles de tous les éléments SIP au fur et à mesure qu'ils sont traversés,
- Le header From indique l'identité de l'entité ayant initié l'enregistrement,
- Le header To indique l'identité de l'UA enregistré,
- Le header Call-ID contient l'identifiant unique qui servira pendant toute la phase d'enregistrement,
- Le header Cseq contient le numéro de séquence d'une phase d'enregistrement. Pour un UA donné, il s'incrémente de 1 à chaque nouvel enregistrement,
- Le header Contact indique l'adresse IP et le port sur lesquels l'UA et le Registrar s'échangent la signalisation SIP,
- Le header Expires indique la durée d'enregistrement (défaut 3600 s), la valeur 0 signifie dé-enregistrement. En cas de désaccord, c'est l'UAS qui fixe la valeur définitive de ce temporisateur. L'UAC renvoie généralement un message Register de « rafraichissement » à la moitié du temps accepté par le serveur.

Le serveur SIP retourne une réponse 200 OK en cas de succès de l'enregistrement. Au préalable la requête «REGISTER » est challengée par le serveur SIP (Réponse 401 ou 407) pour que l'UA chiffre son mot de passe pour authentification.

Requête REGISTER avec authentification



IP publique et port SIP du terminal 4001

Identifiant unique échanges enregistrement courant

Le terminal s'enregistre de nouveau en cryptant le mot de passe à l'aide des informations fournies précédemment par le registrar

Authentification

Afin de ne pas transmettre les informations confidentielles en clair, il est recommandé que les requêtes SIP entre UAS et UAC soient authentifiées. Le process Digest Access Authentication est utilisé pour ce faire.

REGISTER :
L'UAC envoie une requête REGISTER à l'UAS qui répond par un message 401 Unauthorized. WWW-Authenticate contenant notamment l'informations nonce permettant à l'UAC de calculer sa réponse. Le nonce est un identifiant unique qui change à chaque demande d'authentification. On dit que le message REGISTER est « challengé ».
L'UAC émet un nouveau message REGISTER contenant l'entête Authorization dont le champ « response » contient un digest qui résulte d'une fonction de hachage cryptographique mêlant le contenu des champs uri, method, username, password, realm, qop et nonce généralement selon la méthode MD5.
L'UAS compare le champ « response » reçu dans l'entête Authorization avec ce qui est attendu et valide ou pas la demande d'enregistrement en émettant vers l'UAC un message respectivement 200 OK ou 403 Forbidden.

INVITE :
Même principe, sauf que l'UAS peut envoyer un message de réponse 407 Proxy Authentication Required au lieu d'un 401.

Attention : les informations de compte ne sont pas divulguées en clair mais une personne malintentionnée peut essayer de les récupérer.

Authentification REGISTER

Internet Protocol Version 4, Src: 192.168.133.222, Dst: *192.168.133.114*
User Datagram Protocol, Src Port: 9050, Dst Port: 6050
Session Initiation Protocol **(REGISTER)**
  Request-Line: REGISTER sip:192.168.133.114:6050 SIP/2.0
  Message Header
    Allow: INVITE, ACK, CANCEL, BYE, NOTIFY, OPTIONS, REFER, REGISTER, INFO, UPDATE, PRACK
    Call-ID: call-F57C7487-0000-0010-0118-3EC5@192.168.133.222
    [Generated Call-ID: call-F57C7487-0000-0010-0118-3EC5@192.168.133.222]
    Contact: "9999 proxy sip "<sip:9999@192.168.133.222:9050;transport=UDP>;+sip.instance="<urn:uuid:4923fe07-ce03-0072-95e9-3a2c618cdd43>";+org.lin.specs="conference/1.0,ephemeral/1.1,groupchat/1.2,lime"
    Content-Length: 0
    CSeq: 1 REGISTER
    Expires: 90
    From: <sip:9999@192.168.1.20:9050>;tag=c0a80114-2fcf0
    Max-Forwards: 70
    Supported: replaces, outbound, gruu, record-aware
    To: <sip:9999@192.168.133.114:6050>
    User-Agent: SBC 11.0.1v634 UA_Z3T
    Via: SIP/2.0/UDP 192.168.133.222:9050;branch=z9hG4bK-UX-c0a8-0114-4c8d6

Internet Protocol Version 4, Src: *192.168.133.114*, Dst: 192.168.133.222
User Datagram Protocol, Src Port: 6050, Dst Port: 9050
Session Initiation Protocol (401)
  Status-Line: SIP/2.0 **401 Unauthorized**
  Message Header
    Via: SIP/2.0/UDP 192.168.133.222:9050;received=192.168.133.222;branch=z9hG4bK-UX-c0a8-0114-4c8d6
    Call-ID: call-F57C7487-0000-0010-0118-3EC5@192.168.133.222
    [Generated Call-ID: call-F57C7487-0000-0010-0118-3EC5@192.168.133.222]
    From: <sip:9999@192.168.1.20>;tag=c0a80114-2fcf0
    To: <sip:9999@192.168.133.114>;tag=z9hG4bK-UX-c0a8-0114-4c8d6
    CSeq: 1 REGISTER
    WWW-Authenticate: Digest realm="Zebra3Tel",nonce="1725355283/b8eba39120d3bf3d9132a52513037d1d",opaque="4d14ca33137be629",algorithm=md5,qop="auth"
    Server: UAS_Z3T
    Content-Length: 0

Internet Protocol Version 4, Src: 192.168.133.222, Dst: *192.168.133.114*
User Datagram Protocol, Src Port: 9050, Dst Port: 6050
Session Initiation Protocol **(REGISTER)**
  Request-Line: REGISTER sip:192.168.133.114:6050 SIP/2.0
  Message Header
    Allow: INVITE, ACK, CANCEL, BYE, NOTIFY, OPTIONS, REFER, REGISTER, INFO, UPDATE, PRACK
    [truncated]Authorization: Digest realm="Zebra3Tel", nonce="1725355283/b8eba39120d3bf3d9132a52513037d1d",opaque="4d14ca33137be629",algorithm=md5,qop=auth, username="7010", uri="sip:sip1.zebra3.tel", response="13374b6a5f48b5d8aacffa0e
    Call-ID: call-F57C7487-0000-0010-0118-3EC5@192.168.133.222
    [Generated Call-ID: call-F57C7487-0000-0010-0118-3EC5@192.168.133.222]
    Contact: "9999 proxy sip "<sip:9999@192.168.133.222:9050;reg-key;transport=UDP>;+sip.instance="<urn:uuid:4923fe07-ce03-0072-95e9-3a2c618cdd43>";+org.lin.specs="conference/1.0,ephemeral/1.1,groupchat/1.2,lime"
    Content-Length: 0
    CSeq: 2 REGISTER
    Expires: 90
    From: <sip:9999@192.168.1.20:9050>;tag=c0a80114-2fcf0
    Max-Forwards: 70
    Supported: replaces, outbound, gruu, record-aware
    To: <sip:9999@192.168.133.114:6050>
    User-Agent: SBC 11.0.1v634 UA_Z3T
    Via: SIP/2.0/UDP 192.168.133.222:9050;branch=z9hG4bK-UX-c0a8-0114-4c8d7

Internet Protocol Version 4, Src: *192.168.133.114*, Dst: 192.168.133.222
User Datagram Protocol, Src Port: 6050, Dst Port: 9050
Session Initiation Protocol (200)
  Status-Line: SIP/2.0 **200 OK**
  Message Header
    Via: SIP/2.0/UDP 192.168.133.222:9050;received=192.168.133.222;branch=z9hG4bK-UX-c0a8-0114-4c8d7
    Call-ID: call-F57C7487-0000-0010-0118-3EC5@192.168.133.222
    [Generated Call-ID: call-F57C7487-0000-0010-0118-3EC5@192.168.133.222]
    From: <sip:9999@192.168.1.20>;tag=c0a80114-2fcf0
    To: <sip:9999@192.168.133.114>;tag=z9hG4bK-UX-c0a8-0114-4c8d7
    CSeq: 2 REGISTER
    Date: Tue, 03 Sep 2024 09:21:24 GMT
    Contact: <sip:9999@192.168.133.222:9050;transport=UDP;reg-key>;expires=88
    Expires: 90
    Server: UAS_Z3T
    Content-Length: 0

Authentification INVITE

Internet Protocol Version 4, Src: 192.168.133.222, Dst: 192.168.133.114
User Datagram Protocol, Src Port: 9050, Dst Port: 6050
Session Initiation Protocol (**INVITE**)
Request-Line: INVITE sip:6666@192.168.133.114:6050;user=phone SIP/2.0
Message Header
Allow: INVITE, ACK, CANCEL, BYE, NOTIFY, OPTIONS, REFER, REGISTER, INFO, UPDATE, PRACK
Call-ID: call-D7CD7887-0000-0010-0118-3EC7@192.168.133.222
[Generated Call-ID: call-D7CD7887-0000-0010-0118-3EC7@192.168.133.222]
Contact: <sip:9999@192.168.133.222:9050;transport=UDP>
Content-Length: 333
Content-Type: application/sdp
CSeq: 2 INVITE
From: "9999 proxy sip " <sip:9999@192.168.133.222:9050;user=phone>;tag=c0a80114-2fcf6;sgid=3
Max-Forwards: 69
Min-SE: 600
P-Asserted-Identity: "9999 proxy sip " <sip:9999@192.168.133.222:9050;user=phone>
Session-Expires: 3600
Supported: replaces,update,timer,100rel
To: <sip:6666@192.168.133.114:6050;user=phone>
User-Agent: SBC 11.0.1v634 UA_Z3T
Via: SIP/2.0/UDP 192.168.133.222:9050;branch=z9hG4bK-UX-c0a8-0114-4c8e0
X-Z3T-Diagnostics: SBCInternal;cid=11182;media-mode="audio:DSP video:N/A";tdmchannel="b:0 t:3 g:1 c:2"
Message Body

Internet Protocol Version 4, Src: 192.168.133.114, Dst: 192.168.133.222
User Datagram Protocol, Src Port: 6050, Dst Port: 9050
Session Initiation Protocol (401)
Status-Line: SIP/2.0 **401 Unauthorized**
Message Header
Via: SIP/2.0/UDP 192.168.133.222:9050;received=192.168.133.222;branch=z9hG4bK-UX-c0a8-0114-4c8e0
Call-ID: call-D7CD7887-0000-0010-0118-3EC7@192.168.133.222
[Generated Call-ID: call-D7CD7887-0000-0010-0118-3EC7@192.168.133.222]
From: "9999 proxy sip " <sip:9999@192.168.133.222;user=phone>;tag=c0a80114-2fcf6;sgid=3
To: <sip:6666@192.168.133.114;user=phone>;tag=z9hG4bK-UX-c0a8-0114-4c8e0
CSeq: 2 INVITE
WWW-Authenticate: Digest realm="Zebra3Tel",nonce="1725355292/b28366604d85e3eb4e2d5c83a8422993",opaque="636a743b5fa7c3ce",algorithm=md5,qop="auth"
Server: UAS_Z3T
Content-Length: 0

Internet Protocol Version 4, Src: 192.168.133.222, Dst: 192.168.133.114
User Datagram Protocol, Src Port: 9050, Dst Port: 6050
Session Initiation Protocol (**ACK**)
Request-Line: ACK sip:6666@192.168.133.114:6050;user=phone SIP/2.0
Message Header
Allow: INVITE, ACK, CANCEL, BYE, NOTIFY, OPTIONS, REFER, REGISTER, INFO, UPDATE, PRACK
Call-ID: call-D7CD7887-0000-0010-0118-3EC7@192.168.133.222
[Generated Call-ID: call-D7CD7887-0000-0010-0118-3EC7@192.168.133.222]
Contact: <sip:9999@192.168.133.222:9050;transport=UDP>
Content-Length: 0
CSeq: 2 ACK
From: "9999 proxy sip " <sip:9999@192.168.133.222:9050;user=phone>;tag=c0a80114-2fcf6;sgid=3
Max-Forwards: 69
To: <sip:6666@192.168.133.114;user=phone>;tag=z9hG4bK-UX-c0a8-0114-4c8e0
User-Agent: SBC 11.0.1v634 UA_Z3T
Via: SIP/2.0/UDP 192.168.133.222:9050;branch=z9hG4bK-UX-c0a8-0114-4c8e0
X-Z3T-Diagnostics: SBCInternal;cid=11182;media-mode="audio:DSP video:N/A";tdmchannel="b:0 t:3 g:1 c:2"

Internet Protocol Version 4, Src: 192.168.133.222, Dst: 192.168.133.114
User Datagram Protocol, Src Port: 9050, Dst Port: 6050
Session Initiation Protocol (**INVITE**)
Request-Line: INVITE sip:6666@192.168.133.114:6050;user=phone SIP/2.0
Message Header
Allow: INVITE, ACK, CANCEL, BYE, NOTIFY, OPTIONS, REFER, REGISTER, INFO, UPDATE, PRACK
[truncated]Authorization: Digest realm="Zebra3Tel", nonce="1725355292/b28366604d85e3eb4e2d5c83a8422993",opaque="636a743b5fa7c3ce",algorithm=md5,qop=auth, username="7010", uri="sip:6666@sip1.zebra3.tel", response="6fbad03752b20706249
Call-ID: call-D7CD7887-0000-0010-0118-3EC7@192.168.133.222
[Generated Call-ID: call-D7CD7887-0000-0010-0118-3EC7@192.168.133.222]
Contact: <sip:9999@192.168.133.222:9050;transport=UDP>
Content-Length: 333
Content-Type: application/sdp
CSeq: 3 INVITE
From: "9999 proxy sip " <sip:9999@192.168.133.222:9050;user=phone>;tag=c0a80114-2fcf6;sgid=3
Max-Forwards: 69
Min-SE: 600
P-Asserted-Identity: "9999 proxy sip " <sip:9999@192.168.133.222:9050;user=phone>
Session-Expires: 3600
Supported: replaces,update,timer,100rel
To: <sip:6666@192.168.133.114:6050;user=phone>
User-Agent: SBC 11.0.1v634 UA_Z3T
Via: SIP/2.0/UDP 192.168.133.222:9050;branch=z9hG4bK-UX-c0a8-0114-4c8e1
X-Z3T-Diagnostics: SBCInternal;cid=11182;media-mode="audio:DSP video:N/A";tdmchannel="b:0 t:3 g:1 c:2"
Message Body

Paramètres des entêtes WWW-Authenticate et Authorization

- realm

Chaine de caractères unique durant toute la durée de l'authentification. Le realm est commun aux UAC et UAS engagés dans le process d'authentification

- nonce (server nonce) ou cnonce (customer nonce)

Chaine de caractères aléatoire à usage unique générée dans chaque réponse 401 ou 407 valable pour une seule transaction. Un nonce utilisé précédemment ne doit pas être accepté afin de se protéger contre les attaques. Cette valeur est utilisée pour calculer l'empreinte (par défaut MD5) du mot de passe contenue dans la réponse

- opaque

Chaine de caractères générée par le serveur qui doit être retournée intacte par le client dans les entêtes Authorization des messages subséquents pendant toute la durée du process d'authentification

- qop

Ce paramètre qui définit la qualité de la protection peut prendre les 2 valeurs « auth » (authentification seule) ou « auth-int » (authentification avec contrôle de l'intégrité)

- username

Nom de l'utilisateur au sein du realm

- response

Chaine de caractères déterminée par le hachage des données précédentes selon l'algorithme retenu

- stale

Flag indiquant que le nonce fourni dans la requête en provenance de l'UA est périmé

- algorithm

Algorithme utilisé pour produire le digest. Par défaut MD5. D'autres algorithmes comme SHAxxx peuvent être utilisés si supportés

Registrar SIP

Un registrar SIP est un dispositif d'enregistrement qui permet la **mémorisation de la localisation** (AOR) d'un UA SIP. Un fournisseur de services SIP doit être en capacité d'associer une extension (user part) appartenant à son domaine à la SIP URI diffusée par un terminal lors de son enregistrement,

Exemple : le champ « contact » du message Register contient l'URI 4001@92.187.116.109:63951
La fonction registrar du domaine sipdemo.zebra3tel.com (host part indiquée dans la Request-Line - cf trace précédente -) associera l'extension 4001 (user part) à la localisation 92.187.116.109:63951

Attention : l'authentification de l'utilisateur n'est pas de la responsabilité du registrar à proprement parler.

Définition du dialogue SIP

La RFC 361 fournit pour le dialogue SIP la définition suivante :

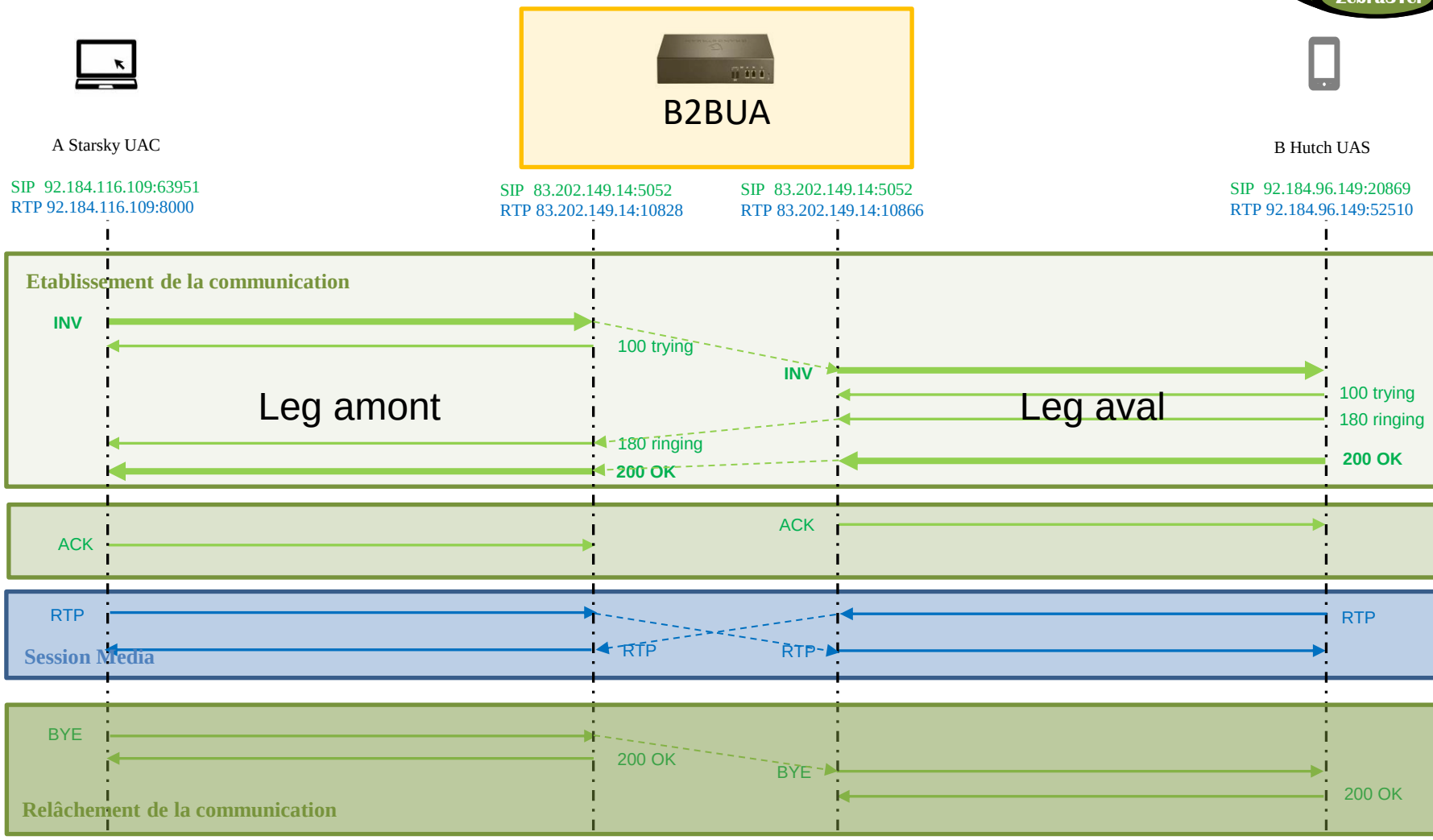
Le dialogue est une relation persistante entre 2 UA SIP. Le dialogue est établi par une réponse de type 200 OK à une requête INVITE. Le dialogue était aussi connu sous l'appellation « call leg » dans les anciennes RFC.
Le dialogue est identifié par la combinaison des tags From et To et Call-ID. La valeur de ces 3 éléments ne varie pas au cours du dialogue.

L'enregistrement n'est pas un dialogue à proprement parler. La séquence de messages qui se produit lors de l'émission d'une requête REGISTER satisfait presque totalement aux critères ci-dessus mais il ne s'agit pas d'un message initial INVITE.

On considère que l'INVITE est le seul message qui initie un dialogue SIP.

Lors de la réception de la réponse 200 OK au message INVITE émis, le dialogue préalablement dans l'état «early dialog» passe à l'état «confirmed dialog».

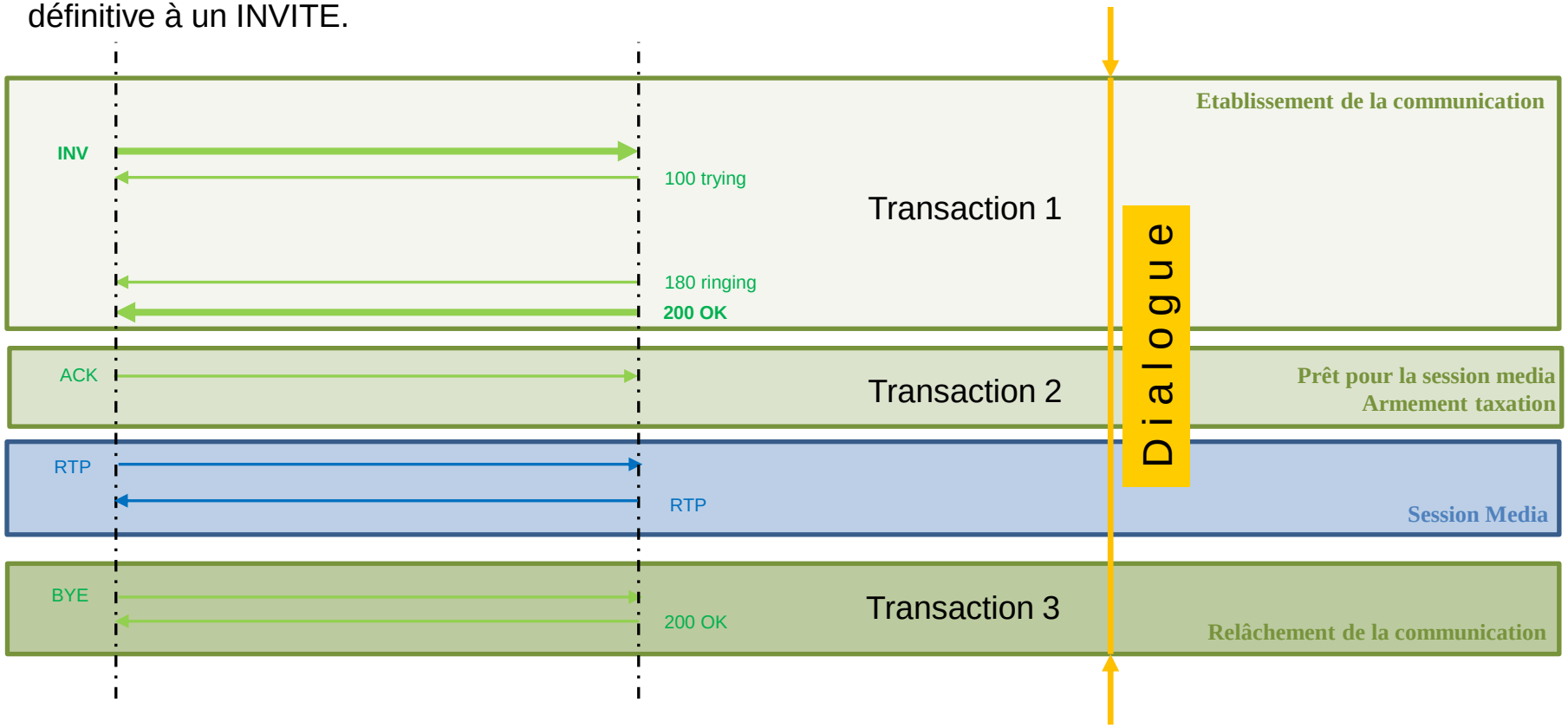
Dialogue SIP de base



En gras : message avec SDP

Dialogue et transactions SIP

Un dialogue SIP entre 2 UA est formé d'au moins 3 transactions elles-mêmes constituées de messages appelés méthodes (commandes ou requêtes) ou réponses qui indiquent soit une information de progression d'appel soit une information d'état final. Une transaction « valide » commence par une commande et se termine par une réponse définitive (> 1xx). Le ACK ci-dessous (ne nécessitant pas de réponse) est une transaction à part entière. Il indique que la session média est établie et confirme la réception d'une réponse définitive à un INVITE.



Un UA considère que la requête ACK lui est destiné si la valeur :

1. du From tag de la requête ACK reçue est égal au From tag qu'il a émis dans la transaction précédente
2. du To tag de la requête ACK reçue est égal au To tag qu'il a émis dans la transaction précédente
3. du Call-ID de la requête ACK reçue est égal au Call-ID qu'il a émis dans la transaction précédente

Illustration par l'exemple, étude d'un cas réel

Call flow

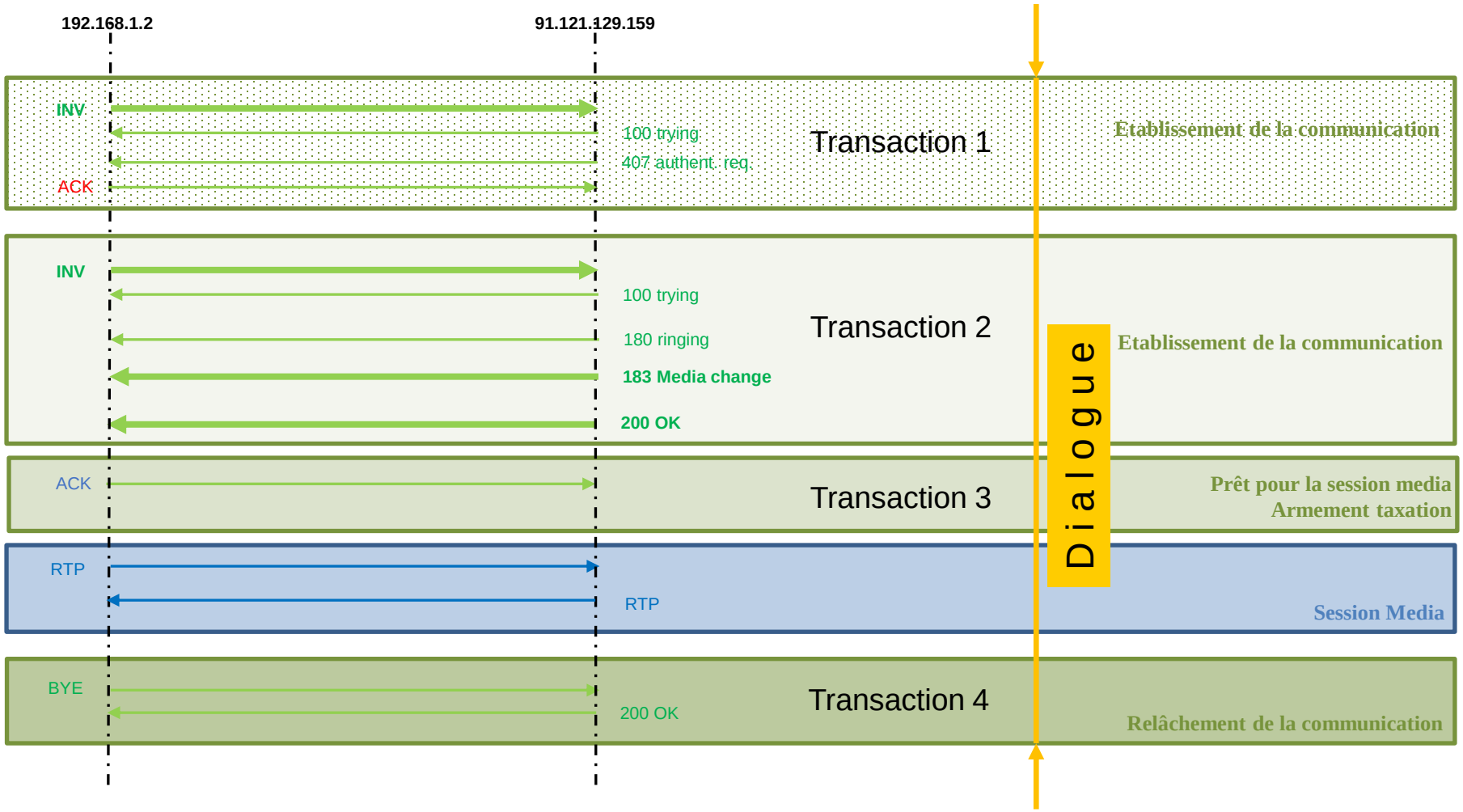


Illustration par l'exemple T1

DIALOGUE

TRANSACTION 1

1-INVITE	src 192.168.1.2	dest 91.121.129.159
	Via branch	z9hG4bKPj368219eb-a636-408d-98f0-01323ed5e1bf
	To tag	
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11450 INVITE
2-100 Trying	src 91.121.129.159	dest 192.168.1.2
	Via branch	z9hG4bKPj368219eb-a636-408d-98f0-01323ed5e1bf
	To tag	
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11450 INVITE
3-407	src 91.121.129.159	dest 192.168.1.2
	Via branch	z9hG4bKPj368219eb-a636-408d-98f0-01323ed5e1bf
	To tag	00-07981-7695c357-734c4be37
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11450 INVITE
4-ACK	src 192.168.1.2	dest 91.121.129.159
	Via branch	z9hG4bKPj368219eb-a636-408d-98f0-01323ed5e1bf
	To tag	00-07981-7695c357-734c4be37
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11450 ACK

Illustration par l'exemple T2

Il s'agit d'un re-INVITE car la session est établie (ACK(407) émis), le Cseq incrémenté, nouvelle transaction mais même dialogue

DIALOGUE

TRANSACTION 2

5-INVITE	src 192.168.1.2	dest 91.121.129.159
	Via branch	z9hG4bKPj43792e53-8bf1-4027-bbe6-6d828968a377
	To tag	
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11451 INVITE
6-100 Trying	src 91.121.129.159	dest 192.168.1.2
	Via branch	z9hG4bKPj43792e53-8bf1-4027-bbe6-6d828968a377
	To tag	
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11451 INVITE
7-180 Ringing	src 91.121.129.159	dest 192.168.1.2
	Via branch	z9hG4bKPj43792e53-8bf1-4027-bbe6-6d828968a377
	To tag	00-07979-7695c35d-55db1dce5
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11451 INVITE
8-183 Media change	src 91.121.129.159	dest 192.168.1.2
	Via branch	z9hG4bKPj43792e53-8bf1-4027-bbe6-6d828968a377
	To tag	00-07979-7695c35d-55db1dce5
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11451 INVITE
9-200 OK	src 91.121.129.159	dest 192.168.1.2
	Via branch	z9hG4bKPj43792e53-8bf1-4027-bbe6-6d828968a377
	To tag	00-07979-7695c35d-55db1dce5
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11451 INVITE

Illustration par l'exemple T3 & 4

DIAL

DIALOGUE

TRAN 3

10-ACK	src 192.168.1.2	dest 91.121.129.159
	Via branch	z9hG4bKPjd05e384b-3d00-4078-a4c0-ea8a2e7b044f
	To tag	00-07979-7695c35d-55db1dce5
	From tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	11451 ACK

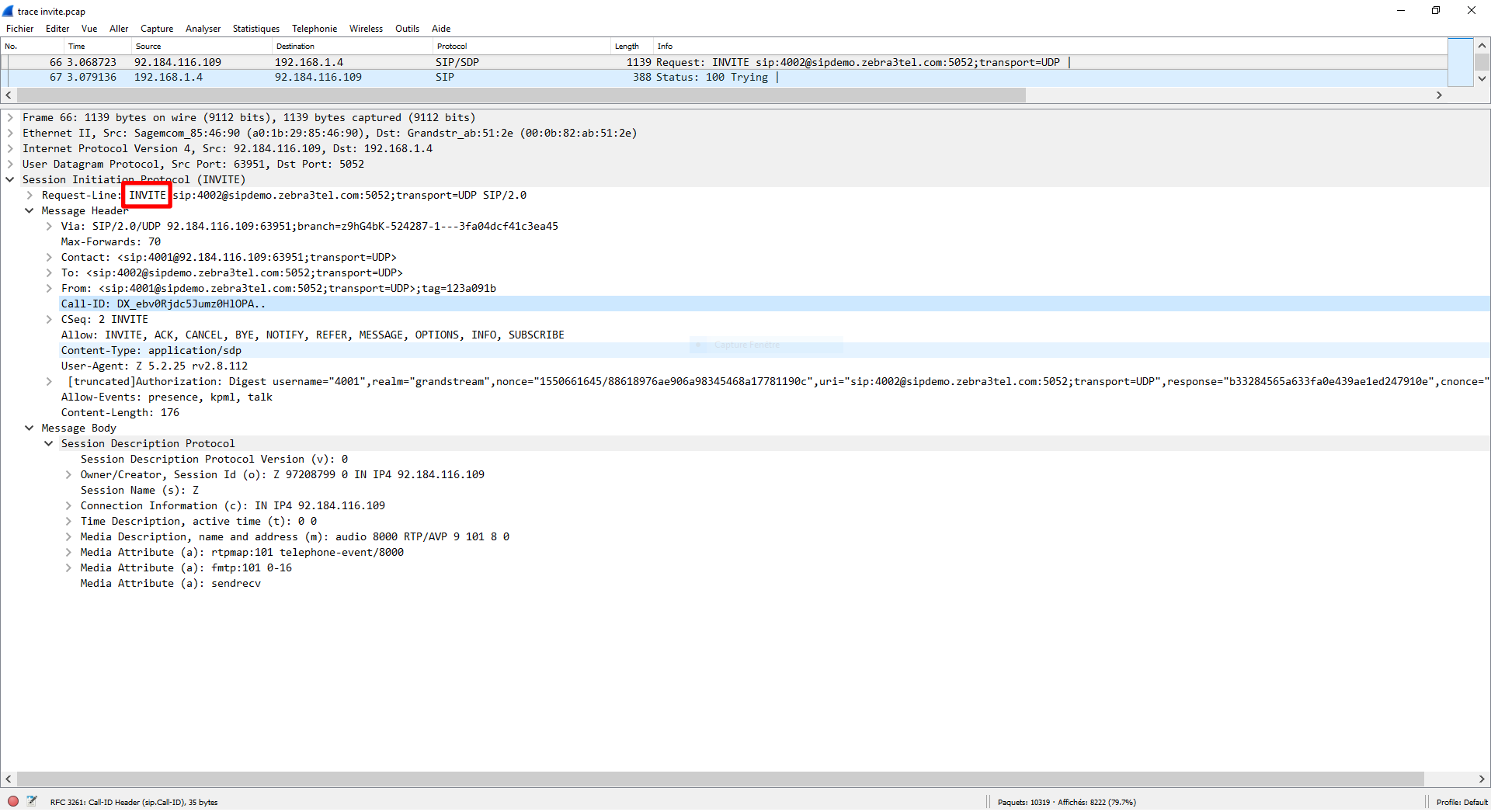
Session Media

TRANSACTION 4

11-BYE	src 91.121.129.159	dest 192.168.1.2
	Via branch	z9hG4bK-AUSQ-59b52c32-71fd2f4a
	To tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	From tag	00-07979-7695c35d-55db1dce5
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	147538974 BYE
12-200 OK	src 192.168.1.2	dest 91.121.129.159
	Via branch	z9hG4bK-AUSQ-59b52c32-71fd2f4a
	To tag	e84183d3-3b25-42c2-9719-b9d25ec8daf8
	From tag	00-07979-7695c35d-55db1dce5
	Call-ID	0dd2ce12-49fb-43f8-8c29-e6aafed9d05b
	CSeq	147538974 BYE

Dans l'exemple ci-dessus, le 1er INVITE est « challengé » pour sécuriser l'échange des données sensibles. L'UA émet un nouvel INVITE appelé re-INVITE puisqu'un message 4-ACK(407) mettant fin à la transaction a été préalablement envoyé.
La méthode 10-ACK(200OK) qui indique que la session média est prête est une transaction à elle seule. Le via branch est constant tout au long de la durée de vie d'une transaction SIP. Le Call-ID est constant durant la durée de vie d'un dialogue. Une fois le message ACK échangé, le dialogue passe à l'état «established dialog».

Trace message SIP INVITE



Adressage SIP

L'adressage en SIP, comme pour toute autre technologie est primordial. C'est en analysant le contenu du « Request-URI » de la Request-Line et en le confrontant à des tables internes que le mécanisme de routage est mis en œuvre. Il s'agit de trouver une voie de sortie pour diriger la demande d'appel vers le nœud réseau suivant. On appelle URI (Uniform Resource Identifier) « l'adresse SIP » permettant d'identifier un utilisateur.

Il existe 2 formats d'URI utilisés en SIP : le tel URI ou le sip URI

Le format sip URI est le plus fréquemment utilisé.

Format des URI

1. sip URI sip:user[:password]@hos[t:port][;uri-parameters?headers]

   user : nom d'utilisateur sous la forme d'une chaine alphanumérique
   password : mot de passe en clair, à ne surtout pas utiliser
   host : serveur hébergeant la ressource SIP, sous forme de FQDN ou IP V4/6, en présence d'un hostname la partie user est obligatoire
   port : port SIP (en l'absence d'indication c'est la valeur 5060 qui est implicite)
   uri-parameters : paramètres additionnels sous la forme nom_param=valeur_param
   Ex :
   sip:4001@sipdemo.zebra3tel.com:9054;transport=UDP;tag=123a091b
   sip:4001@192.168.1.1:5059;transport=UDP
   sip:1001@mycompany.com:15000
   sip:nom.prenom@zebra3tel-voice.com;transport=UDP
   sip:+33184254112@sipdemo.zebra3tel.com;user=phone

   (le paramètre user=phone indique que la partie « user » devra être traitée comme une tel uri)

2. tel URI RFC3966 tel:telephone-subscriber
   C'est un simple numéro de téléphone qui doit être au format global cad sous la forme d'un numéro international au sens E.164. Si tel n'est pas le cas, il doit être représenté sous la forme d'un numéro local avec un descripteur supplémentaire appelé « phone-context »
   Ex :
   tel:+33184254112
   tel:184254112;phone-context=+33
   tel:4001;phone-context=sipdemo.zebra3tel.com:9054

N.B.:
Les formats d'URI tel:4001;phone-context=sipdemo.zebra3tel.com:9054 et sip:4001@sipdemo.zebra3tel.com:9054;transport=UDP ont la même signification.

Adressage de type sip URI

Les sip URI sont les plus couramment utilisées. Leur résolution est un processus en 2 étapes permettant de localiser tout d'abord une entité pour permettre ensuite de joindre l'utilisateur.

Il existe 2 formats :

1. adresse électronique « user@host » comme contact@zebra3.tel. Le host est soit une adresse IP ou bien un nom de domaine pouvant être résolu au travers d'un DNS. C'est l'adressage idéal qu'il est possible de mettre en place dans les réseaux privés permettant la correspondance de la partie « user » et son adresse IP au sein de l'infrastructure privée concernée.
2. adresse téléphonique « numéro-téléphone@host » comme +33184254112@sipdemo.zebra3tel.com:5052. C'est ce qui est le plus couramment utilisé car c'est aujourd'hui le seul moyen d'identifier sans ambiguïté ni erreur un abonné au téléphone grâce au numéro attribué pour son opérateur. La partie host n'est pas exploitée puisque toutes les informations nécessaires sont disponibles dans la partie « user ».

L'utilisation d'URI sous la forme « user@host » reste problématique à ce jour. Il n'y a aucune action concertée entre opérateurs pour définir un service DNS permettant de garantir la résolution fiable de la partie « host ». Les opérateurs n'ont pas non plus pour la plupart d'entre eux, mis en place de base de données « aliasing » pour résoudre la partie « user ».

Il en résulte que l'exploitation de la partie « user » sans traiter la partie « host » du numéro de téléphone au format global international est la seule alternative à l'heure actuelle pour acheminer des demandes d'appel dans les réseaux publics.

Négociation des capacités media

SIP utilise la syntaxe SDP (Session Description Protocol) selon la RFC4566 pour négocier les caractéristiques media d'une communication principalement pendant la phase d'établissement de celle-ci. Les caractéristiques media peuvent aussi être modifiées un fois la communication SIP établie.

La négociation media entre 2 UA SIP est basés sur le modèle « offre/réponse » utilisant le SDP. Chaque entité participant au dialogue échange avec les autres entités ses « facultés » media c'est-à-dire la listes des codecs supportés et leur description ainsi que l'adresse IP et le port media sur lequel il écoutera.
Une description media SDP peut être présente dans le « Message Body » des messages SIP INVITE SIP Re-INVITE, ACK,18x, 200 OK.

La structure SDP comprend des lignes qui décrivent les attributs du media commençant par la ligne m=. Ces propositions sont classées par ordre de préférence. Dans le cas des types de media dynamiques une ligne « a= » est ajoutée pour le décrire.

Le SDP est composé d'une série de lignes ou paramètres appelés descripteurs <caractère>=<valeur> <CR><LF>, où <caractère> est un caractère alphabétique sensible à la casse et <valeur> est un texte dont la structure dépend du type d'attribut.

Le SDP est composé de trois sections descriptives principales :

1. **la session,**
2. **le timing,**
3. **le media,**

Chaque message peut comporter plusieurs descriptions de timing et de media, mais seulement une description de session.
Le SDP se trouve dans la partie « Message Body » des messages INVITE, Re-INVITE,18x, 200 OK, ACK.

Descripteurs du SDP session, timing et media

Description de la session
v= (numéro de version du protocole, actuellement seulement 0)
o= (identificateur d'origine et de session: nom d'utilisateur, identifiant, numéro de version, adresse réseau)
s= (nom de la session: obligatoire avec au moins un caractère codé UTF-8)
i=* (titre de la session ou information courte)
u=* (URI de description)
e=* (zéro ou plus adresse e-mail avec le nom facultatif des contacts)
p=* (zéro ou plus numéro de téléphone avec le nom optionnel des contacts)
c=* (informations de connexion - pas nécessaire si inclus dans tous les media , il faut une ligne c dans le SDP)
b=* (zéro ou plus de lignes d'information de bande passante)

Description du timing (obligatoire)
t= (time the session is active)
r=* (zéro ou plus de répétition)
z= * (ajustements du fuseau horaire)
k=* (clé de chiffrement)
a=* (zéro ou plus de lignes d'attribut de session)

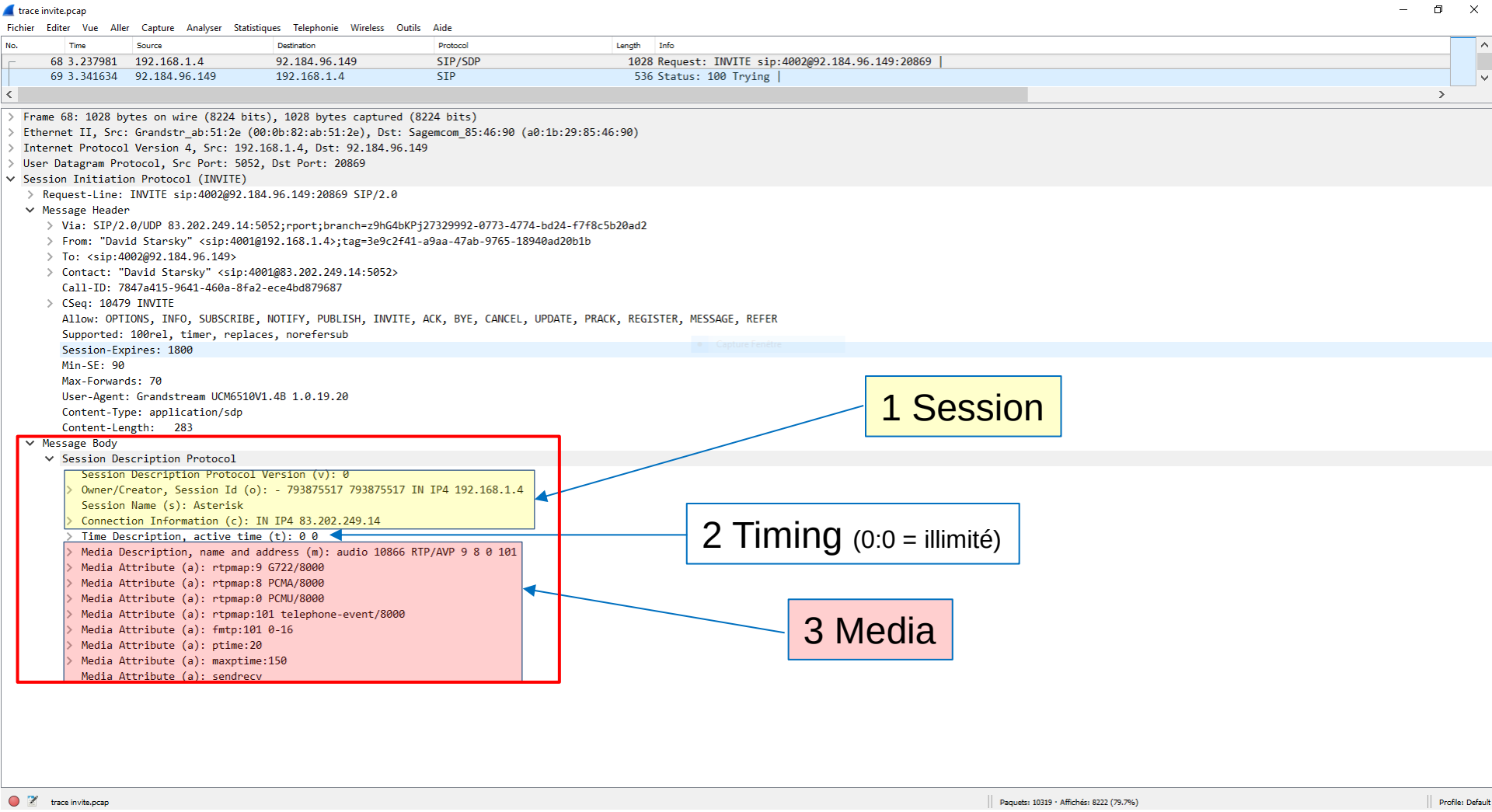
Description du media (si présent)
m= (nom du media et adresse du transport)
i=* (titre du media ou champ d'information)
c=* (informations de connexion - facultatif si inclus au niveau de la session, il faut une ligne c dans le SDP)
b=* (zéro ou plus de lignes d'information de bande passante)
k=* (clé de chiffrement)
a=* (zéro ou plus de lignes d'attribut multimédia - surpassant les lignes d'attributs)

* Ligne optionnelle

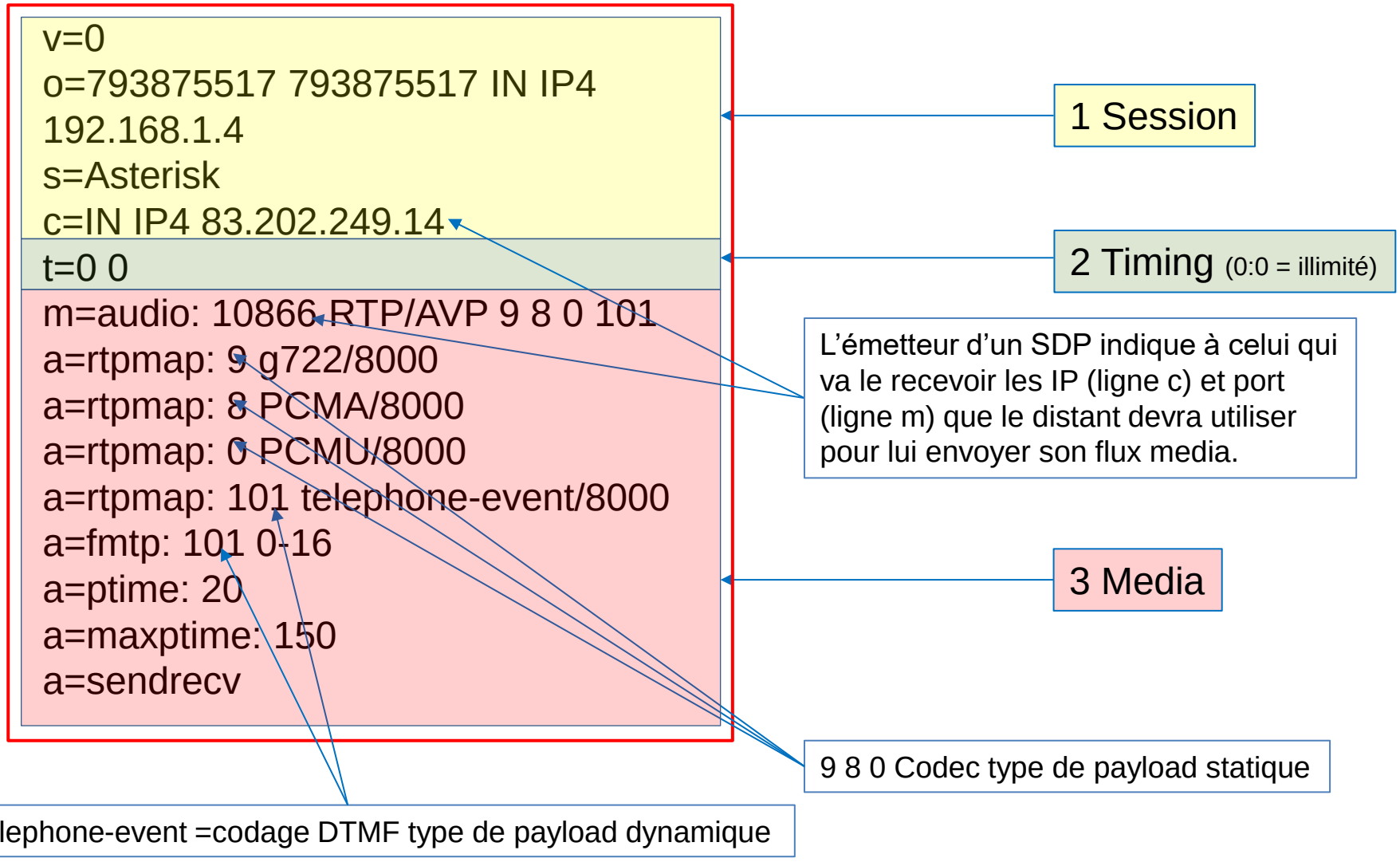
Descripteurs et attributs SDP

Descriptor	Comment
a=	Attributes to extend SDP in the form a=<attribute> or a=<attribute>:<value>.
b=	Contains information about the bandwidth required for the session or media in the form b=<bandwidth_type>:<bandwidth>.
c=	Connection data about the session including the network type (usually IN for Internet), address type (IPv4 or IPv6), the connection source address, and other optional information. For example: c=IN IPv4 10.31.101.20
i=	A text string that contains information about the session. For example: i=A audio presentation about SIP
k=	Can be used to convey encryption keys over a secure and trusted channel. For example: k=clear:444gdduudjffdee
m=	Media information, consisting of one or more lines all starting with m= and containing details about the media including the media type, the destination port or ports used by the media, the protocol used by the media, and a media format description. m=audio 49170 RTP 0 3 m-video 3345/2 udp 34 m-video 2910/2 RTP/AVP 3 56 Multiple media lines are needed if SIP is managing multiple types of media in one session (for example, separate audio and video streams). Multiple ports for a media stream are indicated using a slash. 3345/2 udp means UDP ports 3345 and 3346. Usually RTP uses even-numbered ports for data with the corresponding one-higher odd ports used for the RTCP session belonging to the RTP session. So 2910/2 RTP/AVP means ports 2910 and 2912 are used for RTP and 2911 and 2913 are used for RTCP. Media types include udp for an unspecified protocol that uses UDP, RTP or RTP/AVP for standard RTP and RTP/SAVP for secure RTP.
o=	The sender's username, a session identifier, a session version number, the network type (usually IN for Internet), the address type (for example, IPv4 or IPv6), and the sending device's IP address. The o= field becomes a universal identifier for this version of this session description. For example: o=PhoneA 5462346 332134 IN IP4 10.31.101.20
r=	Repeat times for a session. Used if a session will be repeated at one or more timed intervals. Not normally used for VoIP calls. The times can be in different formats. For example: r=7d 1h 0 25h r=604800 3600 0 90000
s=	Any text that describes the session or s= followed by a space. For example: s=Call from inviter
t=	The start and stop time of the session. Sessions with no time restrictions (most VoIP calls) have a start and stop time of 0. t=0 0
v=	SDP protocol version. The current SDP version is 0 so the v= field is always: v=0
z=	Time zone adjustments. Used for scheduling repeated sessions that span the time between changing from standard to daylight savings time. z=2882844526 -1h 2898848070 0

Position du SDP dans une requête SIP INVITE



Exemple de SDP commenté



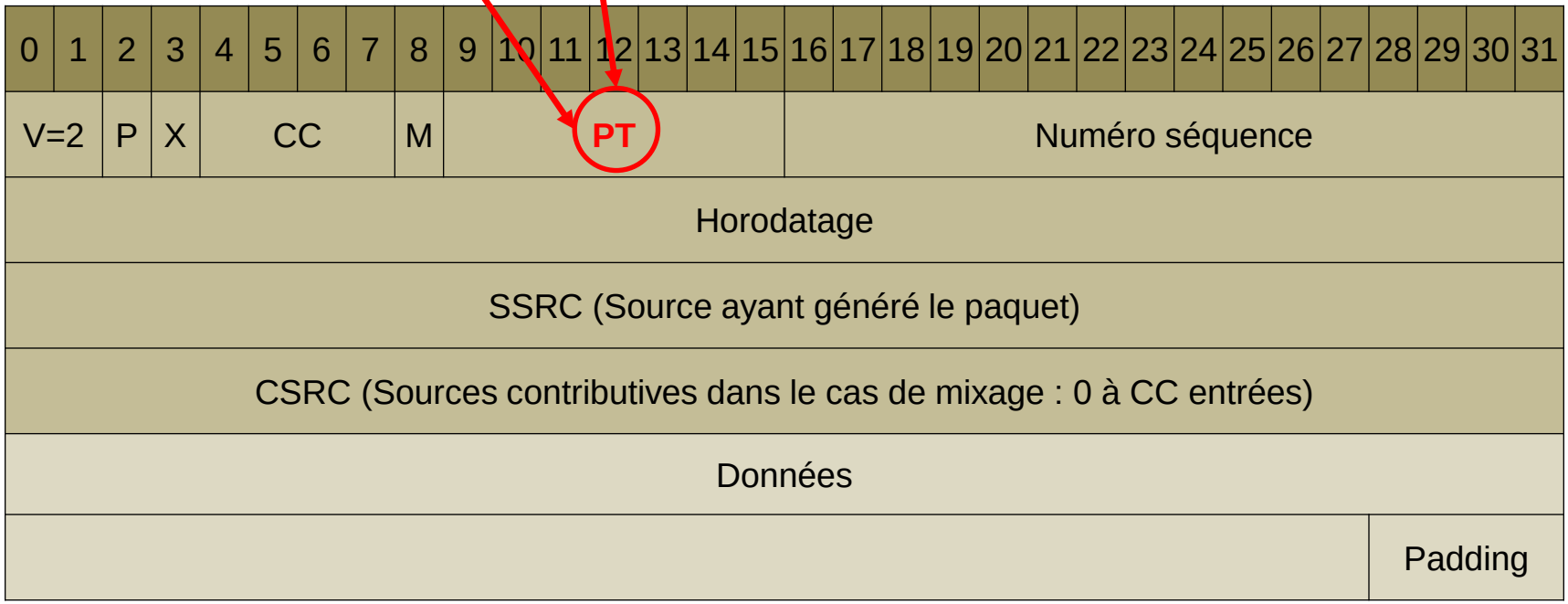
Relation SDP et type de payload RTP

SIP SDP

```
m=audio 10866 RTP/AVP 9 101
a=rtpmap:9 g722/8000
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-16
a=ptime:20
a=maxptime:150
a=sendrecv
```

La valeur du champ « payload type » de l'entête RTP correspond au numéro du type de codec négocié pendant l'établissement de l'appel.
La valeur 9 indique que le flux media est codé en G.722 et la valeur 101 indique que les événements DTMF sont transportés dans le RTP selon format défini dans la RFC4733.

Media RTP



Négociation media

Les extensions portées par les UA ainsi que les trunks VoIP possèdent leurs propres caractéristiques media. Les UAS associent un profil codecs aux extensions qu'ils gèrent et aux trunks VoIP auxquels ils sont raccordés. Ces profils codecs permettent entre autres de limiter les possibilités media devant être gérées, ils agissent comme des filtres «ne laissant passer» que ce qui est autorisé. Les extensions ou les trunks VoIP doivent avoir au moins un codec commun avec les profils codecs auxquels ils sont associés. Si tel n'est pas le cas, la communication ne pourra pas s'établir.

L'offre SDP émise par un UAC comprend dans la m-line une liste ordonnée de codecs. L'UAS réceptrice devra confronter la liste de codecs reçue dans l'offre SDP avec le profil codec (il s'agit également d'une liste ordonnée de codecs) correspondant à l'UAC ayant transmis son offre. S'il existe au moins un codec commun, l'UAS en conserve la trace dans le contexte d'appel sur le leg entrant (amont) puis transmet la liste ordonnée des codecs communs vers le leg sortant (aval) en supprimant les codecs non supportés dans le profil codec du leg entrant. Cette liste de codecs acceptée et filtrée en entrée est ensuite confrontée au profil codecs configuré sur le leg aval pour ne retenir que les codecs autorisés en sortie. Si l'offre SDP n'est pas acceptable (par exemple aucun codec supporté), l'UA envoie une erreur 488.

L'UA aval transmet dans sa **réponse** SDP une liste de codecs correspondant aux caractéristiques media qu'il supporte. Cette liste ordonnée ne peut pas comporter des **codecs différents** de ceux reçues dans l'**offre**. Par contre cette liste peut comprendre moins de codecs et l'ordre de ceux-ci peut différer de celui dans lequel ils avaient été préalablement reçus dans l'offre SDP. **Théoriquement, c'est le premier codec commun entre l'offre et la réponse SDP qui est retenu sur le leg aval.**

L'UAS émet ensuite vers l'UAC amont la liste ordonnée des codecs retenus après traitement sur le leg aval. Cette liste est confrontée au profil codecs configuré sur le leg amont pour filtrage avant son émission. **Théoriquement, c'est le premier codec commun entre l'offre et la réponse SDP qui est retenu sur le leg amont.**

Transcodage media

Lorsque les codecs retenus sur les legs amont et aval sont différents, le media transitant par l'élément réseau doit être adapté. C'est ce qu'on appelle le transcodage. Le transcodage ne doit pas être confondu avec l'encodage du signal étudié plus haut.

Les UA de type IP PBX ou SBC tenteront naturellement de tout mettre en œuvre pour éviter le transcodage en modifiant l'ordre des codecs dans les offres et réponses SDP qu'ils transmettent. Les bonnes pratiques dictent que le codec G.711 devrait toujours être présent dans les configurations des terminaux ou trunks VoiP ainsi que dans les profils codecs qui leur sont associés. On parle de codec G.711 pivot.

Dans certains cas, il est impossible de conserver un codec unique de part et d'autre d'un nœud réseau du fait des configurations codecs dans les terminaux ou les trunks VoIP. Ces cas sont rares mais peuvent exister.

Du fait des traitements supplémentaires nécessaires, un transcodage est destructif au moins sur un voie, il altère souvent la qualité vocale de bout en bout et induit un délai plus ou moins perceptible. Un transcodage n'est pas «visible» dans le SDP lorsqu'il se produit à distance d'un UA. De multiples transcodages sur le chemin emprunté par un communication téléphonique longue distance peuvent se produire.

A noter que, pour un même codec, lorsque les durées de paquetisation (ligne a=ptime:xx) entre legs amont et aval sont différentes, il y a aussi transcodage.

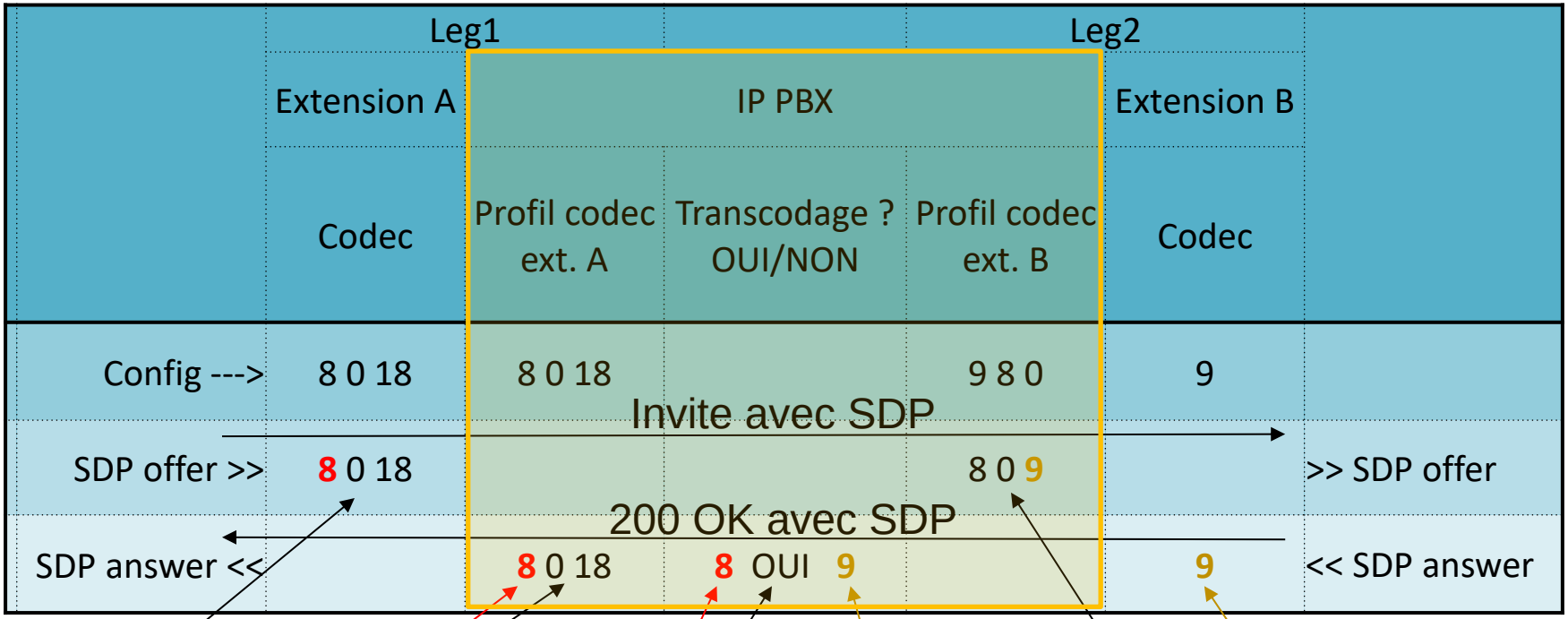
La négociation media a pour but de sélectionner un codec sur les legs amont et aval d'un serveur SIP traitant le media. C'est le protocole SDP qui est utilisé pour négocier les capacités media des terminaux, trunks VoIP ou des éléments réseaux se trouvant sur le chemin d'une communication.
Lorsqu'il est impossible de sélectionner un codec commun sur les legs amont et aval, le media doit être adapté. On parle alors de transcodage souvent mis en œuvre dans des composants électroniques spécifiques appelés DSP.
Le changement temps réel du format de codage d'un flux numérique nécessite une puissance de calcul importante et variable en fonction du couple de codecs concerné. Puisque le signal incident devra d'abord être décodé puis de nouveau encodé, le transcodage est donc une opération néfaste pour la qualité vocale.

Quelques exemples de négociation media

Trace id		Leg1			Leg2		
		Extension A	IP PBX			Extension B	
		Codec	Profil codec ext. A	Transco ?	Profil codec ext. B	Codec	
1	Config --->	9 8 0	9 8 0		8 0 18	8 0 18	
	SDP offer >>	9 8 0			8 0 18		>> SDP offer
	SDP answer <<		8 0 9	8 NON 8		8 0 18	<< SDP answer
4	Config --->	0	9 8 0		8 0 18	8 0 18	
	SDP offer >>	0			0 8 18		>> SDP offer
	SDP answer <<		0	0 NON 0		0 8 18	<< SDP answer
5	Config --->	9	9 8 0		8 0 18	8 0 18	
	SDP offer >>	9			8 0 18		>> SDP offer
	SDP answer <<		9	9 OUI 8		8 0 18	<< SDP answer
6	Config --->	3	9 8 0		8 0 18	8 0 18	
	SDP offer >>	3	INACCEPTABLE 488				>> SDP offer
	SDP answer <<						<< SDP answer
17	Config --->	8 0 18	8 0 18		9 8 0	9	
	SDP offer >>	8 0 18			8 0 9		>> SDP offer
	SDP answer <<		8 0 18	8 OUI 9		9	<< SDP answer
23	Config --->	8 0 18	8 0 18		9 8 0	123 2 97 3 18 0 8 9	
	SDP offer >>	8 0 18			8 0 9		>> SDP offer
	SDP answer <<		8 0 18	8 NON 8		8 0 9	<< SDP answer
25	Config --->	iLBC (dyn)	iLBC (dyn)		Opus (dyn) 9	Opus (dyn) 9	
	SDP offer >>	97			9		>> SDP offer
	SDP answer <<		97	97 OUI 9		9	<< SDP answer
26	Config --->	iLBC (dyn)	iLBC (dyn)		Opus 9 iLBC	Opus 9 iLBC	
	SDP offer >>	97			97 9		>> SDP offer
	SDP answer <<		97	97 NON 97		97 9	<< SDP answer
30	Config --->	9 Opus iLBC	Opus iLBC 9		iLBC Opus 9	Opus iLBC	
	SDP offer >>	9 123 97			9 123 97		>> SDP offer
	SDP answer <<		123 97 9	123 NON 123		123 97	<< SDP answer
33	Config --->	9	iLBC Opus 9		iLBC Opus 9	iLBC	
	SDP offer >>	9			9 97		>> SDP offer
	SDP answer <<		9	9 OUI 97		97	<< SDP answer

Implémentation UCM version 1.0.20.23. Transcodage OPUS impossible --> 488

Exemple commenté négociation media



1) Offre SDP de A vers l'IP PBX

2) Offre codecs réordonnée proposée par l'IP PBX vers B en fonction de ce qui a été reçu sur le leg amont

3) Réponse SDP de B vers l'IP PBX
Codec retenu par B

4) Codec retenu leg aval

5) Codec retenu leg amont
Les codecs négociés en amont et aval étant différents, il y a transcodage

6) Réponse SDP de l'IP PBX vers A
Codec retenu par A

Détermination des informations de connexion

La RFC 3261 indique que le récepteur d'une requête SIP doit router ses réponses en utilisant les IP ou FQDN et port SIP présents dans l'entête **Via** le plus récent.

Par ailleurs l'entête **Contact** du Message Header indique les IP ou FQDN et port SIP de l'émetteur (A) qui doivent être utilisés par le récepteur (B) afin que (B) puisse adresser (A) lorsqu'il lui envoie des requêtes subséquentes (cad dans le même dialogue SIP).

La RFC 4566 (puis 8866) dit que les lignes (c) **Connexion Information** et (m) **Media Description** du SDP indiquent les IP et port media de l'émetteur (A) qui doivent être utilisés par le récepteur (B) afin que (B) puisse adresser (A) lorsqu'il lui envoie son flux media.

```
Internet Protocol Version 4, Src: 86.246.146.226, Dst: 192.168.1.3
User Datagram Protocol, Src Port: 52382, Dst Port: 5070
Session Initiation Protocol (INVITE)
  Request-Line: INVITE sip:5000@sip1.zebra3.tel SIP/2.0
  Message Header
    Via: SIP/2.0/UDP 192.168.4.136:52382;branch=z9hG4bKy2UE2Vh1G5Rj;rport
    From: "1001 UCM1" <sip:1001@sip1.zebra3.tel>;tag=f4YhCOI3GE5u
    To: "5000 440HD-1-L1 DIV-3E" <sip:5000@sip1.zebra3.tel>
    Call-ID: KU-pXxKJhT0m4UrxpxWh
    [Generated Call-ID: KU-pXxKJhT0m4UrxpxWh]
    CSeq: 427 INVITE
    Contact: "1001 UCM1" <sip:1001@86.246.146.226:52382;x-reg=6B88C98B97720D31>;audio
      SIP C-URI display info: "1001 UCM1"
      Contact URI: sip:1001@86.246.146.226:52382;x-reg=6B88C98B97720D31
      Contact parameter: audio
    Content-Type: application/sdp
    Content-Length: 266
    User-Agent: Sipnetic/1.1.4 Android
    Supported: 100rel,timer,replaces,tdialog
    Allow: INVITE,ACK,OPTIONS,BYE,CANCEL,UPDATE,INFO,SUBSCRIBE,NOTIFY,REFER,PRACK,MESSAGE
    Session-Expires: 300
    Max-Forwards: 70
  Message Body
    Session Description Protocol
      Session Description Protocol Version (v): 0
      Owner/Creator, Session Id (o): - 1839308223 1839308223 IN IP4 86.246.146.226
      Session Name (s): -
      Connection Information (c): IN IP4 86.246.146.226
      Time Description, active time (t): 0 0
      Media Description, name and address (m): audio 64202 RTP/AVP 9 8 0 101
```

Cohabitation SIP / NAT

Tout se passe sans trop de problème lorsque les UA SIP sont directement connectés à Internet, ce qui est rarement le cas. La plupart du temps, les UA SIP sont localisés derrière des fonctionnalités NAT.
L'IETF a défini différents types de NAT (a-full cone, b-restricted cone, c-port-restricted cone, d-symmetric). Ces notions sont vagues et on préfère parler d'association (1-endpoint independant, 2-endpoint dependant, 3-address and port dependant). Généralement l'association se base sur l'adresse et le port d'émission privés sans prendre en compte la destination (cas 1-)

Il est fondamental que les UA « récupèrent » leur IP publiques afin de rester « joignables de l'extérieur ». Plusieurs techniques sont disponibles (RPORT, STUN, TURN, ICE…) pour contourner ce problème. Ces techniques ne sont pas toutes intégrées dans les équipements SIP. Les SBC gèrent très bien cette problématique ... ce n'est pas le cas de tous les terminaux et IPPBX.
Il faut être très vigilant lorsqu'on déploie des solutions VoIP, chaque industriel ayant sa propre implémentation pour contourner cette difficulté.

Les problèmes de NAT se traduisent :

- au niveau SIP par une impossibilité d'établir une communication ou de s'enregistrer,
- au niveau media par une perte du flux RTP dans 1 voire les 2 sens (one-way audio, dead air call).

Cohabitation SIP / NAT

rport : ce paramètre (vide ou pas) si présent dans l'entête Via des requêtes permet à son récepteur de router ses réponses en utilisant les IP et port sources présents dans l'entête IP plutôt que les IP et port contenus dans header Via reçu comme indiqué par la RFC 3261. La RFC 3581 décrit ce mécanisme de liaison (private / public binding) qui modifie la RFC 3261.

La réponse inclut les paramètres « rport » et « received » dans le header via valorisés respectivement aux port et IP source reçus dans l'entête IP de la requête.
Cela permet aux serveurs SIP d'associer les IP et port « publics » aux IP et port « privés » reçus dans le Via de la requête et d'envoyer leurs réponses en utilisant les IP et port « publics » pour joindre l'émetteur.

Tout se passe sans trop de problème lorsque les UA SIP sont directement connectés à Internet, ce qui est rarement le cas. La plupart du temps, les UA SIP sont localisés derrière des routeurs NAT.

STUN : ce protocole est défini par la RFC 3489. Il permet à un client UDP de découvrir ses IP et port publics ainsi que le type de NAT derrière lequel il se trouve. Le client SIP utilise les informations publiques récupérées pour valoriser les champs **Contact** du Message Header, **Connexion Information** et **Media Description** du SDP. Si ces champs sont valorisés avec des IP privées le flux RTP sera impossible à établir.

4. Infrastructures

Entités fonctionnelles SIP

Différentes fonctions sont impliquées dans la gestion des extensions SIP et dans le cheminement d'une communication SIP :

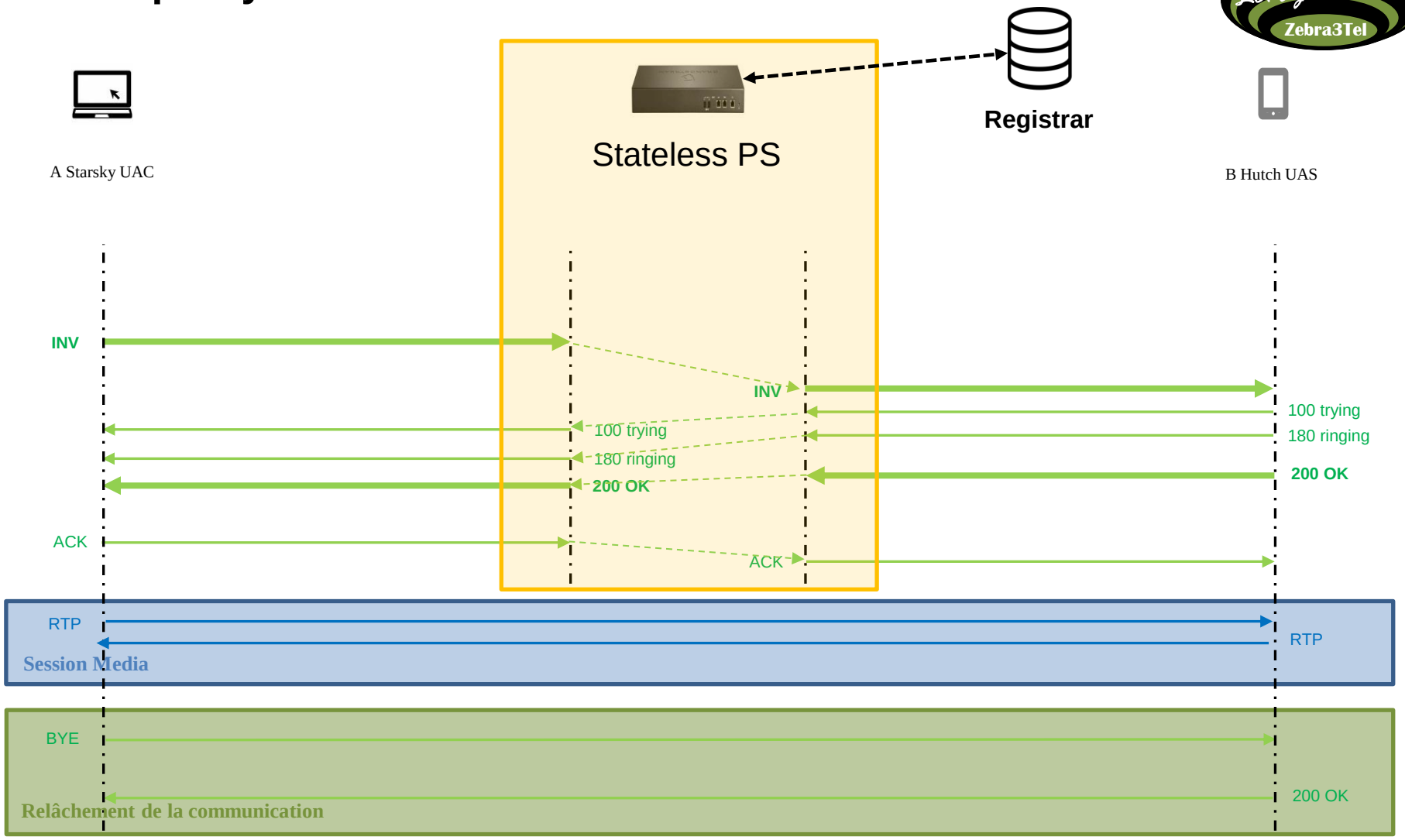
1. Le **proxy server** qui peut (stateful) ou peut ne pas (stateless) contrôler complètement l'état de l'appel. Le proxy est une entité qui se comporte comme un client d'un côté et comme un serveur de l'autre en relayant sans trop les modifier les messages lui parvenant. Le proxy server devrait être transparent aux messages qui le traversent. A l'instar des proxy HTTP, le proxy SIP permet de valider les requêtes qu'il reçoit, d'authentifier les utilisateurs, de résoudre des noms de domaine, …
   - Un stateful proxy conserve l'état de l'appel en gérant les transactions. Ils sont notamment capables de réaliser des fonctions de routages évolués et peuvent générer des CDR (ticket de taxation) ou des rapports statistiques,
   - Un stateless proxy ne fait que traiter l'info qui convient pour choisir la destination suivante. Il se contente de relayer les messages tels qu'il les a reçus. En particulier tous les messages relatifs à une communication SIP contiennent le même call-ID. Ce type de proxy ne se met jamais dans le « chemin » media. **Le stateless proxy est le seul type de serveur SIP qu'on peut considérer comme étant transparent vis-à-vis du contenu des messages SIP le traversant**.
2. Le **registrar** pour enregistrer les terminaux,

Entités fonctionnelles SIP

3. **Le back-to-back user agent (B2BUA) est un stateful proxy qui regénère systématiquement les messages qu'il renvoie. Il n'est absolument pas transparent et introduit ses propres entêtes. C'est cette implémentation qu'on retrouve dans la plupart des réseaux SIP. Un B2BUA peut embarquer de très nombreuses fonctionnalités (comme le transcodage media) et délivrer des services évolués par lui-même ou bien en se connectant à d'autres plateformes pour y trouver des informations supplémentaires (serveur Enum pour la portabilité par exemple). Le B2BUA se met généralement dans le chemin media en mentionnant sa propre IP dans les offres et réponses SDP qu'il génère.**

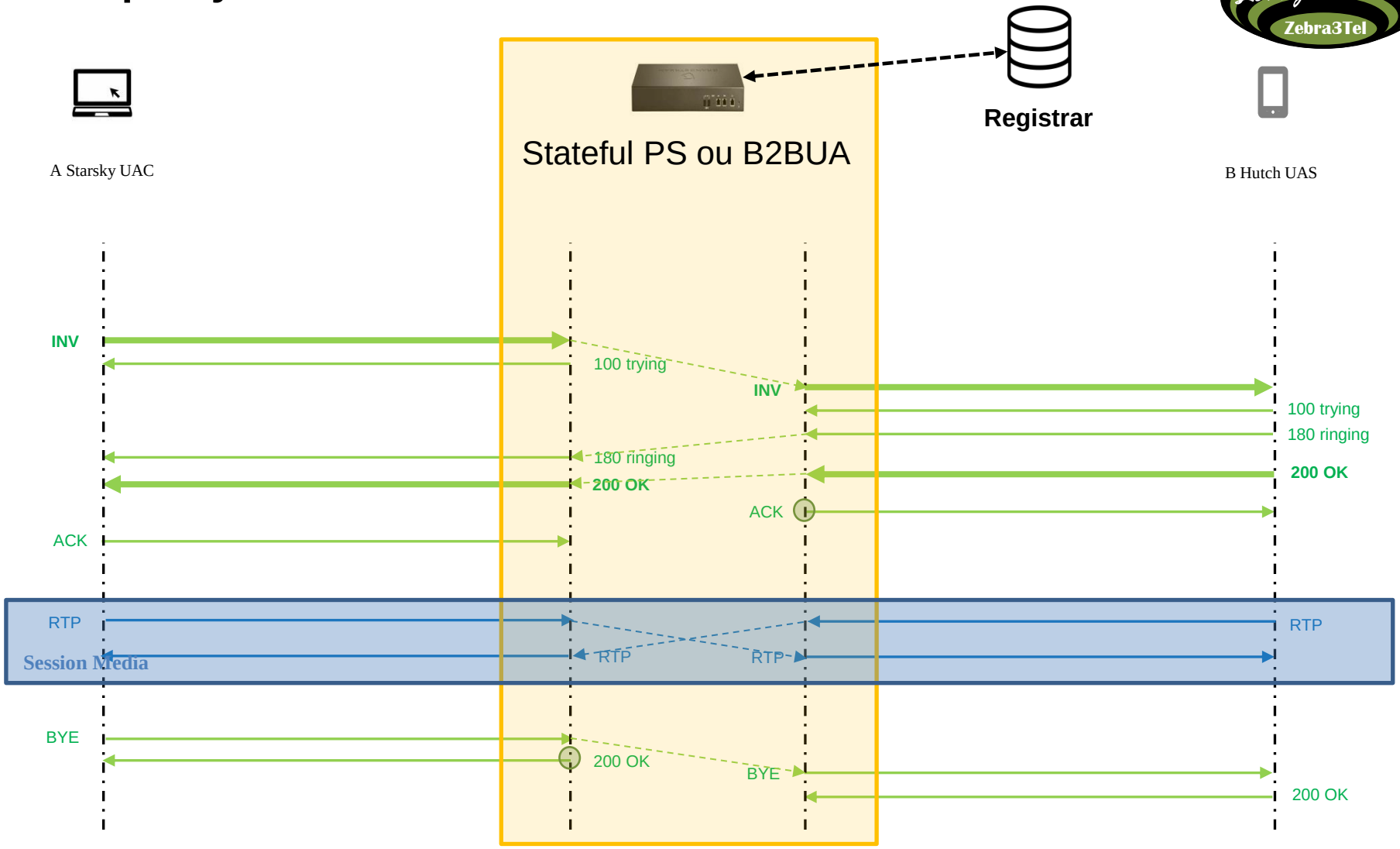
4. Le **redirect server (RS)** répond à une requête invite par un message 3xx ou bien rejette l'appel. Une réponse 3xx peut par exemple renvoyer des informations de localisation ou de routage concernant l'URL présente dans la requête initiale. Ces informations se trouvent dans le champ « contact » de la réponse. Le Redirect Server permet « d'alléger » la charge de l'entité qui le contacte puisque le RS « embarque » une partie de l'intelligence de routage comme le ferait un serveur applicatif.

Stateless proxy server



Gestion de l'établissement d'appel seulement

Stateful proxy server ou B2BUA



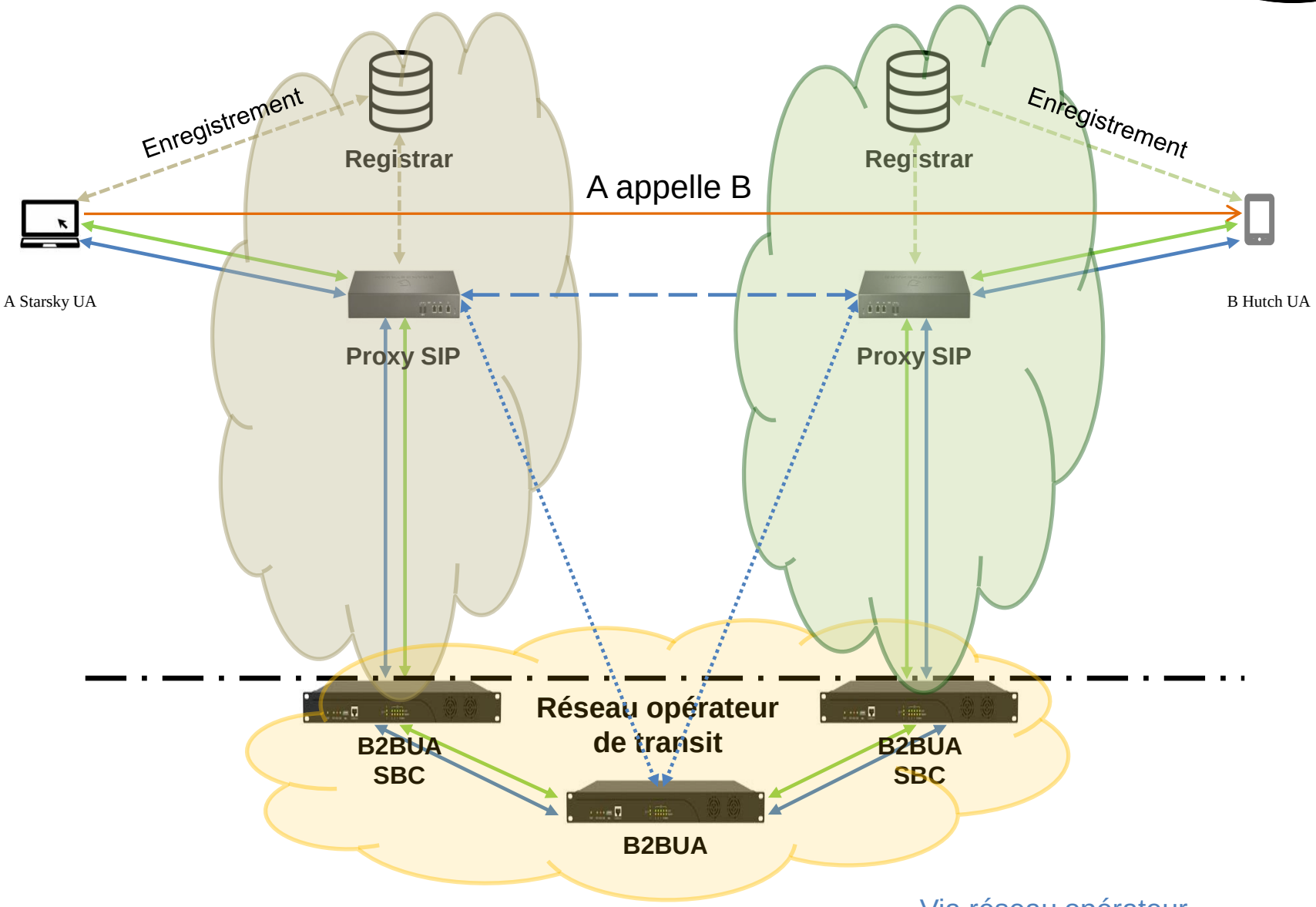
Le stateful proxy server ou B2BUA :

- peut s'interfacer au Registrar,
- gère tous les flux signalisation,
- est capable d'adapter le media,

Exemples de stateful proxy server ou B2BUA : SBC, IP PBX

○ Message généré par le B2BUA

Cheminement d'une communication SIP



CDR

Tous les proxy SIP doivent produire des CDR (Call Detail Record) dès qu'ils reçoivent un message INVITE. Ces-ci sont plus ou moins détaillés et contiennent à minima les champs suivants :

- date et heure de début de la communication,
- durée de la communication lorsqu'elle est efficace,
- numéro appelant,
- numéro appelé,
- route entrante,
- route sortante.

Ces CDR sont des enregistrements qui :

- contiennent toutes les informations nécessaires pour la facturation client,
- servent à établir des rapports de trafic à des fins statistiques,
- apportent un des informations très utiles lors de sessions de troubleshooting,
- doivent-être produites aux autorités compétentes en cas de réquisition.

En plus de ces informations les proxy SIP peuvent générer des traces d'appel, logs, alarmes, journaux et statistiques plus ou moins précis qui sont utilisés par les exploitants.

Aspect Réglementaire

En France la violation du secret de la correspondance, qu'elle circule par voie postale ou par un réseau de télécommunication, est réprimée par les articles 226-15 et 432-9 du code pénal et par l'article L33-1 du code des postes et communications électroniques.

Art. 226-15 Le fait, commis de mauvaise foi, d'ouvrir, de supprimer, de retarder ou de détourner des correspondances arrivées ou non à destination et adressées à des tiers, ou d'en prendre frauduleusement connaissance, est puni d'un an d'emprisonnement et de 45 000 euros d'amende. Est puni des mêmes peines le fait, commis de mauvaise foi, d'intercepter, de détourner, d'utiliser ou de divulguer des correspondances émises, transmises ou reçues par la voie des télécommunications ou de procéder à l'installation d'appareils conçus pour réaliser de telles interceptions.

Sujets non traités dans ce cours

1. Transmission des fax protocole T.38,
2. Dimensionnement et supervision réseau,
3. Traduction : analyse, routage, acheminement,
4. Troubleshooting SIP,
5. Qualité de service, RTCP, files d'attente, gigue,
6. Qualité vocale, les codecs, phénomène d'écho,
7. Sécurité, DOS et TDOS, exposition à l'internet, SBC
8. Détection et lutte contre la fraude, spam

Glossaire

3GPP	3rd Generation Partnership Project	
ABS	Analysis By Synthesis	
ADPCM	Adaptive Differential Pulse Code Modulation	
ADSL	Asymmetric Digital Subscriber Line	
ATA	Adaptateur de Terminal Analogique	
B2BUA	Back to Back User Agent	
CAN	Convertisseur Analogique Numérique	ADC
CDR	Call Detail Record	
CELP	Code-Excited Linear Prediction	
CNA	Convertisseur Numérique Analogique	DAC
CNG	Comfort Noise Generation	
Codec	Coder / Decoder	
CSM	Cable Sous Marin	
DSP	Digital Signal Processor	
DTMF	Dual Tone Multi Frequency	FV
DTX	Dual Tone Multi Frequency	FV
FH	Discountinuous Transmission	
FO	Fibre Optique	
FV	Fréquence Vocale	
GPRS	General Packet Radio Service	
GSM	Global System for Mobile	
GW	GateWay	
IANA	Internet Assigned Numbers Authority	
IETF	Internet Engineering Task Force	
IP	Internet Protocol	
ISDN	Integrated Services Digital Network	RNIS
LAN	Local Area Network	
MGW	Media Gateway	
MIC	Modulation à Impulsions Codés	PCM

Glossaire

MMUSIC	Multiparty Multimedia Session Control	
MOS	Mean Opinion Score	
OTT	Other The Top	
PABX	Private Automatic Branch eXchange	
PAM	Pulse Amplitude Modulation	
PCM	Pulse Coded Modulation	MIC
PLC	Packet Loss Concealment	
PLMN	Public Land Mobile Network	
PSTN	Public Switches Telephony Network	
QOS	Quality Of Service	
RNIS	Réseau Numérique à Intégration de Services	
RTC	Réseau Téléphonique Commuté	PSTN
RTCP	Real Time Control Protocol	
RTP	Real Time Protocol	
SBC	Session Border Controller	
SDP	Session Description Protocol	
SDP	Session Description Protocol	
SIP	Session Initiation Protocol	
SNR	Signal to Noise Ratio	
TCP	Transmission Control Protocol	
TDM	Time Division Multiplexing	
UAC	User Agent Client	
UAS	User Agent Server	
UDP	User Datagram Protocol	
UMTS	Universal Mobile Telecommunications System	
URI	Uniform Resource Identifier	
URL	Uniform Resource Locator	
VAD	Voice Activity Detection	
WAP	Wireless Application Protocol	

Let's get started

Zebr 3 • Tel

Communiquer différemment

http://zebra3.tel

Visitez Zebra3.tel

Visitez Zebra3.tel