rostom
communication hub

TOUT CE QUE VOUS DEVEZ SAVOIR SUR LE SIP

Powered by
be ip

La popularité du protocole SIP a fait apparaître de nouvelles difficultés et de nouveaux risques, c'est pourquoi, aujourd'hui, Damien Sandras a lancé son équipe Be IP dans la conception d'un « hub de communication » qui, à la façon d'un canif suisse, apporte une réponse appropriée à chaque problématique que pose le SIP.

Dans ce livret, il partage avec nous quelques-unes de ses compétences afin de permettre de mieux comprendre ce qu'est ce fameux protocole « SIP ». Présenté d'une manière claire et structurée, il apporte au lecteur une idée précise des mécanismes du protocole, des intérêts de son adoption et des points d'attention qu'il requiert.

Bonne lecture,

L'équipe Be IP

Table des matières

Session Initiation Protocol

SIP est un protocole standard ouvert de télécommunications multimédia qui a été normalisé et standardisé par l'IETF (The Internet Engineering Task Force). Il a été défini dans sa version 2 dans le RFC (Request for Comments) 3261 datant de 2002. Dans le monde de la voix sur IP, ce protocole définit comment établir un appel, le terminer, renvoyer des codes d'erreur, etc. Il est devenu progressivement la norme du secteur des télécoms pour les communications multimédia, remplaçant peu à peu le H.323 qui posait, entre autres, des problèmes de sécurité.

SIP a été conçu de manière suffisamment générique afin de permettre d'initier différents types de sessions en temps réel, qu'il s'agisse d'initier un appel, des sessions de messagerie instantanée ou encore une conférence audio ou vidéo multicast.

Il faut noter que certaines fonctionnalités supplémentaires, qui ne sont pas relatives à l'établissement de sessions, font partie de RFC additionnels. Par ailleurs, SIP ne réinvente pas la roue et se base sur la réutilisation de protocoles classiques d'Internet tels que DNS, UDP, TCP, ou encore RTP pour le transport des flux audio et vidéo et SDP pour la description des codecs supportés.

Mécanisme d'un appel SIP

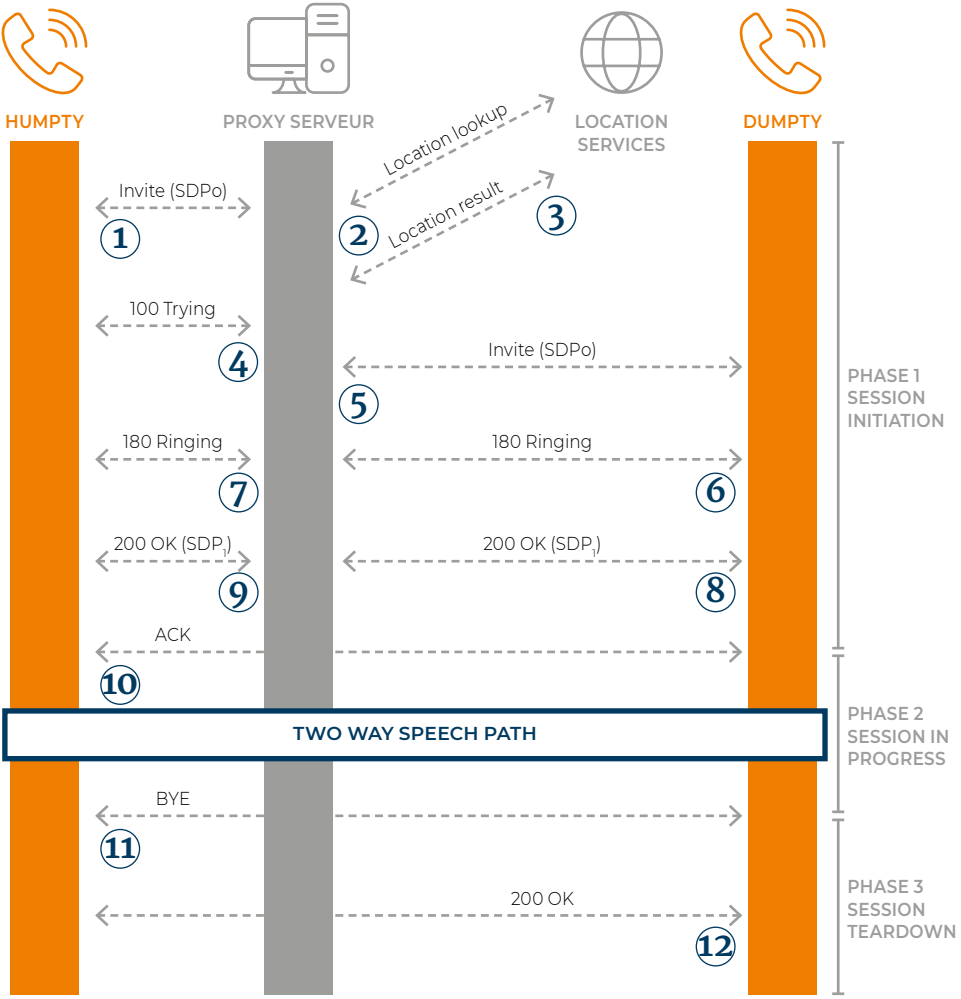
Le protocole SIP, qu'il permette à des téléphones de communiquer entre eux ou encore de se connecter à un fournisseur de minutes SIP, sera identique et repose sur deux types de messages dont certains sont inspirés du fameux protocole HTTP.

LES REQUÊTES

Un UA (User Agent) enverra une requête INVITE afin, par exemple, d'établir un appel. Les requêtes définies par le RFC 3261 sont au nombre de 6. D'autres RFC définissent d'autres requêtes ayant d'autres objectifs, mais tout en conservant les mêmes règles de construction et de routage que celles édictées par le RFC 3261. Cela permet donc à un Proxy Server n'ayant pas la connaissance de ces requêtes de les router d'un UA à un autre sans perte de fonctionnalité.

LES RÉPONSES

Ce sont des réponses à une requête reçue, générées par l'entité distante. Elles sont divisées en 6 classes. Par exemple, une requête d'appel envoyée à un utilisateur SIP inconnu générera une réponse de type « 404 Not Found », réponse que l'on retrouvera également au niveau HTTP.


HUMPTY
PROXY SERVEUR
LOCATION SERVICES
DUMPTY
Location lookup
Location result
Invite (SDPo)
1
2
3
100 Trying
4
Invite (SDPo)
5
180 Ringing
7
180 Ringing
6
200 OK (SDP1)
9
200 OK (SDP1)
8
ACK
10
TWO WAY SPEECH PATH
BYE
11
200 OK
12
PHASE 1 SESSION INITIATION
PHASE 2 SESSION IN PROGRESS
PHASE 3 SESSION TEARDOWN

Exemple de flux d'appel sur base d'une architecture articulée autour d'un simple Proxy Server. Les étapes sont simples :

1. Humpty appelle Dumpty et envoie une requête INVITE au Proxy Server.
2. Le Proxy Server regarde au sein du Location Service quelle est l'adresse IP réelle du poste appelé.
3. Le Proxy Server reçoit la réponse du Location Service, et génère une réponse «100 Trying» indiquant qu'il va tenter de joindre le poste distant.
4. Le Proxy Server relaie l'INVITE vers le poste distant maintenant qu'il connaît sa localisation exacte.
5. Le poste distant accepte l'appel entrant et se met à sonner. Il en informe le proxy via une réponse «180 Ringing».
6. Le Proxy Server relaie la réponse à l'appelant.
7. Le poste distant est décroché par l'utilisateur, l'appel est pris. Il génère une réponse «200 OK».
8. Le Proxy Server relaie la réponse «200 OK» au poste appelant.
9. Le poste appelant envoie une requête «ACK» directement au poste appelé (elle peut passer par le serveur ou pas, suivant la négociation préalable et la configuration).
10. Le flux audio est établi directement entre les deux postes.
11. Quand l'appelant raccroche, il envoie une requête «BYE» au poste appelé (qui peut également passer par le Proxy Server ou pas).
12. La réception de la requête «BYE» est confirmée par l'émission d'un message «200 OK» par le poste qui la reçoit.

Il est à noter que le flux audio ne passera jamais par le Proxy Server car un Proxy Server ne gère que les messages SIP (la signalisation) et pas les flux audio. Il peut en aller autrement dans le cas d'un appel émis à travers un Back-to-Back User.

L'architecture SIP et ses composants

SIP définit donc les composants d'une infrastructure et leurs interactions. Les composants principaux sont les suivants :

User Agent (UA) : de type « Client » ou « Serveur » selon qu'il émet ou reçoit des requêtes SIP. Il s'agit en général des terminaux SIP tels que softphones (téléphone sous forme de logiciel PC, tablette, smartphone), téléphones et certains types de serveurs.

Redirect Server : la composante d'un serveur ou d'un user agent qui permet de rediriger un appel d'un point A à un point B.

Proxy Server : composant principal d'un réseau SIP. Il agit comme un interprète pour les demandes des clients qui recherchent des ressources d'autres serveurs. Il surveille et facilite les échanges.

Le concept a été imaginé pour :

- Être le plus « agnostique » possible afin de supporter de manière transparente les extensions de SIP.
- Conserver aussi peu d'informations que nécessaire afin de contrôler plus finement l'utilisation des ressources mémoire et CPU et ainsi être facilement évolutif.

Registrar : il s'agit d'un autre concept fondamental de tout réseau SIP. Ce composant traite un type de requête SIP spécifique et permettra d'associer à une URI SIP (URI : chaîne de caractères identifiant une ressource sur un réseau) l'emplacement spécifique du terminal sur le réseau. L'association entre URI SIP et adresse IP se fera au sein d'un « Location Service » également défini dans le RFC et qui consiste en une base de données virtuelle.

Back-to-Back User Agent (B2BUA) : entité logique qui reçoit des requêtes SIP et les traite comme le ferait un user agent en mode serveur. Ensuite, pour déterminer comment répondre à la requête, il se comportera comme user agent en mode client. Dit plus simplement, chaque appel entrant génère automatiquement un appel sortant. C'est ainsi que les infrastructures SIP reposant sur un B2BUA auront un contrôle plus fin sur les appels avec des possibilités additionnelles de transcodage des flux audio et vidéo et de manipulation de ces derniers (annonces de pré-décrochage, temporisation…). Si le contrôle est plus fin, l'évolutivité sera pourtant bien moins importante que pour une infrastructure reposant sur un Proxy Server.

On retrouvera donc sur le marché de nombreux serveurs SIP construits autour d'un Proxy Server, et de nombreux serveurs SIP reposant sur un back-to-back user agent. Certains même combinant les deux approches afin d'en tirer le meilleur parti.

Compression et codecs

Le terme codec vient de COder DECoder ou encore de COmpresser ou DECompresser.

Un codec est donc le logiciel ou le matériel qui met en oeuvre un procédé capable de compresser ou décompresser des données dans un format normalisé.

Chaque codec est un compromis entre :

- La qualité de la voix
- La puissance de calcul
- La bande passante
- La latence

Il existe des codecs pour l'audio, mais aussi pour la vidéo.

EN DÉTAILS

La latence doit rester minimale dans le cadre d'un appel en voix sur IP. Elle correspond au décalage entre le temps écoulé entre l'émission de la parole et son écoute par le correspondant. Elle est la somme des différents délais introduits lors de la transmission :

- Capture par le micro du correspondant.
- Conversion en signal numérique.
- Compression par le codec
- Encapsulation en paquets RTP.
- Transmission sur le réseau.
- Somme des délais pour réaliser les opérations inverses à l'autre bout.

LES CODECS AUDIO LES PLUS RÉPANDUS SONT G.711, G.722, G.729 ET OPUS :

- La norme G.711 est la base du transport de la voix sur le réseau téléphonique commuté (RTC, PSTN en anglais) ou sur l'ISDN. Ce codec est également utilisé pour le transport de la voix avec peu de compression dans les réseaux IP, plus généralement sur un LAN et rarement sur Internet à cause de la bande passante nécessaire. Le support de G.711 est rendu obligatoire par la norme WebRTC.

- La norme mondiale de codage G.722 normalisée par l'UIT-T en 1987 permet d'obtenir en voix sur IP une qualité de voix « haute définition » (dite téléphonie large-bande). Cette qualité est obtenue par doublement de la bande de fréquence codée (50-7.000 Hz) par rapport à la qualité téléphonique usuelle dite bande étroite (300-3.400 Hz) produite par le format de codage G.711 utilisé en téléphonie « classique » sur les réseaux RTC. L'utilisateur bénéficie donc d'une sensation de présence de son interlocuteur, d'un confort d'écoute et d'une intelligibilité fortement améliorés.

De nombreux téléphones IP, bien que labellisés « HD Voice » ne font que supporter ce type de signalisation sans pour autant pouvoir le restituer en haute définition réelle. La norme WebRTC ne rend pas obligatoire le support de G.722.

- Le codec G.729 est moins consommateur en bande passante que G.711. Il est utilisé pour obtenir une téléphonie de qualité acceptable à moindre bande passante.
- Le codec OPUS est un format ouvert de compression audio avec pertes, libre de redevances et normalisé par l'Internet Engineering Task Force (IETF) dans le RFC 6716, conçu pour encoder efficacement la voix et plus largement l'audio dans un format unique, tout en ayant une latence suffisamment faible pour la communication en temps réel et une complexité suffisamment faible pour les processeurs embarqués peu puissants. Il supporte divers algorithmes de compression, et peut également utiliser différents algorithmes pour le même fichier audio puisque l'encodeur peut choisir la bitrate, la bande passante, l'algorithme et d'autre détails pour la compression de chaque trame audio. Il est le codec de choix dans la norme WebRTC et commence à être implémenté par certaines marques de téléphones.

De nombreux autres codecs existent, tant pour la voix que pour la vidéo.

Le tableau suivant indique les différents codecs VoIP les plus répandus, ainsi que leur bande passante sur un réseau Ethernet, et leur score MOS.

Le score MOS, obtenu par expérimentation, ou « note d'opinion moyenne » (Mean Opinion Score) est une note donnée à un codec audio pour caractériser la qualité de la restitution sonore. La note varie entre 1 et 5 (excellent).

VOIP-SIP.ORG Codec and Bit Rate

VOIP-SIP.ORG Codec and Bit Rate	Sample Size (Bytes)	Sample Rate (ms)	MOS Quality	Voice Payload Size (Bytes)	Voice Payload Size (ms)
G.711 (64 Kbps)	80	10	4.3	160	20
G.729 (8 Kbps)	10	10	3.7	20	20
G.723.1 (6.3 Kbps)	24	30	3.9	24	30
G.723.1 (5.3 Kbps)	20	30	3.8	20	30
G.726 (32 Kbps)	20	5	3.85	80	20
G.726 (24 Kbps)	15	5	–	60	20
G.728 (16 Kbps)	10	5	3.61	60	30
G.722 (64 Kbps)	80	10	4.13	160	20
iLBC (15.2 Kbps)	38	20	4.14	38	20
iLBC (13.33 Kbps)	50	30	–	50	30

Il est donc nécessaire, si on souhaite conserver une conversation de qualité, de minimiser la latence introduite par la compression/décompression, tout en minimisant la bande passante, la puissance de calcul, mais en conservant la meilleure qualité audio possible.

Signalisation et flux des communications

Packets Per Second (PPS)	Brandwidth Ethernet (Kbps)
50	87.2
50	31.2
33.3	21.9
33.3	20.8
50	55.2
50	47.2
33.3	31.5
50	87.2
50	38.4
33.3	28.8

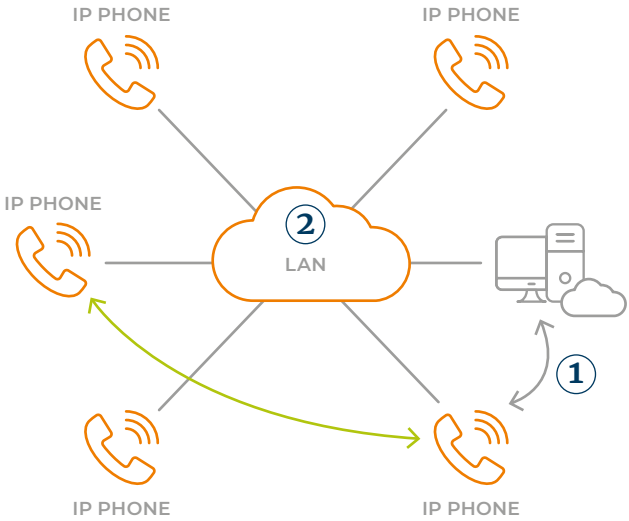
La signalisation échange les informations concernant l'appel, tels que les identifiants des postes, l'état, l'initialisation ou la libération de l'appel.

Qu'il s'agisse d'une architecture sur site ou dans le « Private Cloud », il est important de comprendre le cheminement des flux de communication. La signalisation en SIP utilise très peu de bande passante, alors que le transport des flux est consommateur.

Comme décrit précédemment, il existe plusieurs types d'architectures SIP. Parmi celles-ci, relevons les architectures basées sur un B2BUA, celles qui reposent sur un proxy pur qui ne gèrent que la signalisation ou encore celles s'appuyant sur un mélange des deux composants précédents.

Dans le cas d'un proxy, les flux de média transiteront directement en mode point-à-point, d'un terminal à un autre ou d'un terminal à une passerelle.

Dans le cas d'un B2BUA, les flux de média transiteront par le serveur. Cependant, certaines solutions disposent de mécanismes permettant de basculer les flux directement et dès que possible entre les deux terminaux impliqués dans l'appel.



Ce schéma illustre une architecture typique où les flux ne remontent pas par le serveur dans le Cloud, même en cas d'appel entre deux postes de sites différents. Lors d'un appel sortant vers un numéro public, et dans le cas de l'utilisation d'un SIP trunk, l'appel remontera par contre par l'infrastructure serveur.

En cas de solution multi-sites, il n'est pas rare que les bandes passantes disponibles et réservées à la voix diffèrent d'un site à l'autre. Il est dès lors important que la solution IP dispose d'un mécanisme de Call Admission Control afin de garder le contrôle sur les flux média au sein de votre architecture. Chaque site aura une sortie réseau limitée en bande passante. Le lien privé vers le data center hébergeant le serveur de communications sera également limité. Il est donc nécessaire que la solution de communications garde le contrôle sur les flux.

Le mécanisme de Call Admission Control aura dès lors deux rôles importants pour une qualité de communication irréprochable :

- La compression des appels transitant en inter-sites et vers le DataCenter afin d'économiser la bande passante.
- Le comptage des appels au niveau de chaque site. Tout appel qui pourrait potentiellement surcharger la bande passante allouée garantie sera refusé avec, par exemple, une tonalité de congestion, tout comme cela serait le cas lorsqu'on essaye de faire passer plus d'appels sur une ligne traditionnelle que ce qu'elle ne permet de transporter.

SIP providers

Migrer vers la VoIP et utiliser le protocole SIP pour établir des sessions d'appel permet en toute logique de se passer de la connectivité traditionnelle, tels les accès ISDN et analogiques.

Vous utilisez dès lors la connexion vers Internet afin de faire transiter les messages et requêtes SIP jusqu'à un serveur idéalement placé, qui ressortira sur le réseau classique à un tarif plus avantageux puisque l'appel aura transité par le réseau informatique avant de ressortir. De plus, il est possible d'obtenir des numéros «traditionnels» indépendamment de la localisation géographique de l'entreprise.

Des fournisseurs d'accès permettent d'établir et de recevoir des appels vers et/ou à partir de numéros «traditionnels» en utilisant le protocole SIP. On les appelle ITSP (Internet Telephony Service Provider ou Fournisseur de Services de Téléphonie sur Internet). Connus aussi sous le nom de fournisseur VoIP, ils proposent des SIP trunks à cet effet.

SIP Trunk

Si on prend l'exemple d'une société belge appelant fréquemment aux États-Unis, on peut imaginer l'appel transiter via SIP jusqu'à un serveur SIP situé aux États-Unis et ressortant sur les lignes locales directement là-bas. Ceci permettant donc d'appeler au tarif local américain plutôt qu'au tarif international belge. Cet exemple est bien entendu volontairement simplifié.

PLUS DE POSSIBILITÉS ET DE FLEXIBILITÉ

Le remplacement d'une connexion ISDN au profit d'une connexion SIP comporte quelques avantages comme profiter d'une plus grande flexibilité et de capacités étendues.

- Réduction des coûts de connectivité.
- Une bande passante plus large peut être mise à disposition pour intégrer à la fois voix, vidéo et données.
- La limitation à 2 ou 30 canaux par connexion est abolie au profit d'un dimensionnement plus précis et plus souple.

EN DÉTAILS

Il existe en pratique deux types de fournisseurs de SIP trunking :

- Les opérateurs qui permettent un accès à leur infrastructure à partir de n'importe où dans le monde. Ils ne peuvent toutefois pas garantir la qualité de la communication voix. En effet, les paquets SIP et RTP transiteront par un certain nombre de routeurs IP et de lignes sur lesquels le fournisseur n'a pas de contrôle. En outre, la qualité de ces lignes pourrait être aléatoire en fonction des pannes, du trajet réellement emprunté par les paquets IP et de la charge du réseau.
- Les opérateurs qui offrent un accès à leur infrastructure à partir de leurs liens. Ils contrôlent la qualité du transport de la voix depuis votre entreprise jusqu'à leur infrastructure SIP. Il leur incombe par la suite d'obtenir les bons accords afin de pouvoir acheminer avec qualité la voix, de la manière la moins coûteuse depuis leur infrastructure jusqu'à la destination réelle de l'appel.

Session Border Controller

Un Session Border Controller (SBC) est un équipement hardware ou une application software qui gère et contrôle la façon dont les communications téléphoniques SIP sont initiées, acheminées et clôturées dans un réseau Voice over Internet Protocol (VoIP). Dans la littérature professionnelle, les communications téléphoniques sont habituellement désignées comme des sessions.

SESSION
Communication entre deux parties ou l'appel téléphonique

BORDER
Point de démarcation entre 2 réseaux différents

CONTROL
Capacité à manipuler les flux de données

Les fonctionnalités typiques d'un SBC varient d'un fabriquant à l'autre.

Un SBC fonctionne comme une sorte de pare-feu entre deux réseaux, où seules les sessions autorisées peuvent transiter via le point de raccordement (la frontière).

Parallèlement, le SBC peut assurer l'adaptation des flux échangés par l'intermédiaire du protocole SIP et donc l'interopérabilité. Par la normalisation et la médiation des flux SIP, on peut adopter une connectivité multi-vendeurs et multi-protocoles.

De plus, le SBC définit et contrôle le niveau de la Qualité de Service (QoS) pour toutes les sessions, de telle sorte que les éventuels problèmes de qualité au niveau des communications peuvent être identifiés, et donc plus facilement résolus.

Finalement, le SBC peut en général également agir sur le routage des appels, de telle sorte que les appels d'urgence, par exemple, arrivent directement au bon endroit et reçoivent en outre la priorité sur les autres appels.

Pourquoi un SBC est-il nécessaire ?

Les SBC sont généralement implémentés comme des SIP Back-to-Back User Agents (B2BUA). Un Back-to-Back User Agent opère entre les deux extrémités d'une communication, divise le canal de communication en deux « call legs » et effectue la médiation de toutes les signalisations SIP entre les deux extrémités de l'appel, depuis son initiation jusqu'à la fin, en respectant un routage et une configuration prédéfinie.

FLEXIBILITÉ

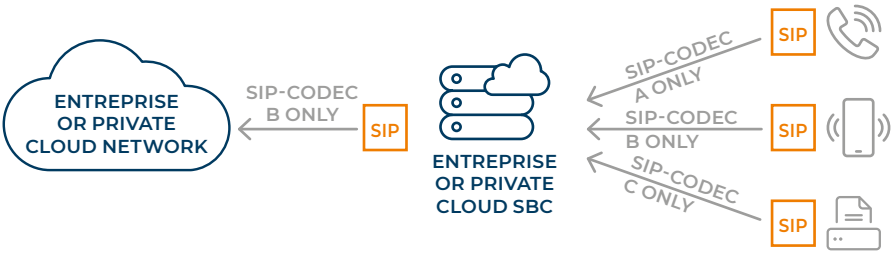
En séparant son réseau du réseau du fournisseur de services, l'utilisateur possède une totale indépendance. Ainsi, des changements peuvent être rapidement effectués sans devoir préalablement négocier avec le fournisseur de services.

Les flux de signalisation et le RTP peuvent être normalisés entre les différents réseaux en interne et en externe, complétés par les adaptations nécessaires au niveau du protocole, le transcodage média et le contrôle de la qualité des communications.

De plus, différents SIP trunks, provenant de différents fournisseurs de réseaux et/ou de systèmes VoIP, peuvent être connectés et routés de manière appropriée par l'intermédiaire du SBC.

De plus en plus de SIP trunks venant de différents fournisseurs de réseaux peuvent être reliés au Cloud privé.

Dans des cas exceptionnels, un soutien peut même être offert pour des flux d'appel qui ne sont pas, en temps normal, pris en charge par le réseau du fournisseur d'accès.



SÉCURITÉ

- Le SBC offre une protection contre les vulnérabilités du protocole SIP.
- Une protection également contre les attaques DoS (Denial-of- Service/par déni de service) en provenance du réseau public.

LOGGING & REPORTING

- Enregistrement de chaque statut d'appel, QoS et monitorage SLA.
- Enregistrement des tentatives d'intrusion.
- Enregistrement des sessions.
- Facturation.

Un SBC placé du côté de l'entreprise est généralement considéré comme un E-SBC (Entreprise Session Border Controller). Le SBC fonctionne ici essentiellement comme un dispositif de sécurité et normalise les connexions SIP des appareils et solutions SIP internes vis-à-vis du côté opérateur public.

Mon opérateur m'offre la protection de son SBC, est-ce suffisamment efficace ?

Confier la sécurité des flux à son seul opérateur, c'est en faire à la fois juge et partie en cas de défaillance ou d'attaque externe.

L'apport d'un élément tiers dans votre chaîne de sécurité permet de consolider l'étanchéité de votre réseau, mais de surcroît, en cas de faille, d'éviter de vous retrouver dans l'inconfortable situation où il est de votre responsabilité de démontrer l'origine de la panne.

Un SBC opérateur a pour fonction essentielle de protéger l'opérateur lui-même plutôt que chaque client individuellement.

Pouvez-vous être certain de la fiabilité de l'ensemble des utilisateurs de cette ressource commune ?

Web Real-Time Communication

L'idée du WebRTC est de faire de tout appareil connecté un dispositif de communication et pouvoir partager en temps réel un maximum d'informations (appels, messages instantanés, partage de fichiers, partage d'écran) entre navigateurs WEB.

Les applications WebRTC offrent par exemple la possibilité d'appeler en ligne un conseiller d'un magasin dont nous sommes en train de consulter le site en cliquant sur le bouton « appel », faire appel à de l'assistance en ligne ou créer un réseau social dynamique interne à l'entreprise.

La richesse et les fonctionnalités dépendront de la fonction que l'on voudra donner à cette interface.

Avec les Communications Unifiées, c'est aussi l'avènement du numéro unique. L'approche WebRTC permettrait de se passer à la fois de l'installation d'applications, des circuits traditionnels des opérateurs téléphoniques et même à terme, du numéro de téléphone.

En 2017, la spécification WebRTC 1.0 passe de l'ébauche de travail au stade plus abouti de « recommandation candidate ».

Ce nouveau standard s'avère en tout cas riche en possibilités et offrira aux entreprises de nombreuses nouvelles applications créatives.

Gosip - Communication platform for human beings – Powered by Be IP.

Damien Sandras
Ingénieur civil
en informatique.

Il y a 20 ans, Damien Sandras écrivait les premières lignes de son logiciel de voix sur IP et de visioconférence ; GnomeMeeting deviendra Ekiga, qui n'est autre que le logiciel de communication de la plupart des distributions GNU/Linux.

Destiné à favoriser les plateformes de communication basées sur des standards ouverts tels que SIP, ce qui n'était au départ qu'une thèse d'étude finira par se retrouver un peu partout : à bord des sous-marins nucléaires de l'armée américaine, il fait une apparition dans la série Mr Robot et même dans la solution PC Linux proposée par NeufCegetel !

Dans la foulée, Damien Sandras devient l'un des membres fondateurs du FOSDEM qui offre aujourd'hui à plus de 8.000 développeurs du monde Open Source l'occasion de se rencontrer, de partager des idées et de collaborer.

C'est en 2004 qu'il part de zéro pour développer un tout nouvel outil de communication IP destiné à bousculer les centraux téléphoniques traditionnels. Aujourd'hui Gosip compte plus de 40.000 utilisateurs.

rostom
communication hub

Be IP
Parc scientifique Fleming
Fond Jean Pâques, 4
B-1348 Louvain-La-Neuve
+32 10 60 87 87

WWW.ROSTOM.TECH