

Travail de Bachelor

Chiffrement / Signature d'emails
Non Confidentiel

Étudiant :	Bonjour Mickael
Travail proposé par :	Duc Alexandre Nom de l'entreprise/institution Adresse NPA Ville
Enseignant responsable :	Duc Alexandre
Année académique	2019-2020

Yverdon-les-Bains, le 7 juillet 2020

Mickael Bonjour

Département TIC
Filière Télécommunications
Orientation Sécurité de l'information
Étudiant Bonjour Mickael
Enseignant responsable Duc Alexandre

Travail de Bachelor 2019-2020

Chiffrement / Signature d'emails

Nom de l'entreprise/institution

Résumé publiable

Dans ce travail de bachelor je vais analyser les solutions actuelles de messagerie sécurisée afin d'identifier les primitives cryptographiques dont j'ai besoin. Ensuite j'établirais un *Proof Of Concept* amenant la technologie choisie dans un cadre de messagerie électronique sécurisée.

Étudiant :	Date et lieu :	Signature :
Bonjour Mickael
Enseignant responsable :	Date et lieu :	Signature :
Duc Alexandre
Nom de l'entreprise/institution :	Date et lieu :	Signature :
Nom Prénom de la personne confiant l'étude

Préambule

Ce travail de Bachelor (ci-après TB) est réalisé en fin de cursus d'études, en vue de l'obtention du titre de Bachelor of Science HES-SO en **Ingénierie / Economie d'entreprise**.

En tant que travail académique, son contenu, sans préjuger de sa valeur, n'engage ni la responsabilité de l'auteur, ni celles du jury du travail de Bachelor et de l'Ecole.

Toute utilisation, même partielle, de ce TB doit être faite dans le respect du droit d'auteur.

HEIG-VD

Vincent Peiris
Chef de département TIC

Yverdon-les-Bains, le 7 juillet 2020

Authentication

Le soussigné, Mickael Bonjour, atteste par la présente avoir réalisé ce travail et n'avoir utilisé aucune autre source que celles expressément mentionnées.

Yverdon-les-bains, le 7 juillet 2020

Mickael Bonjour

Table des matières

1	Cahier des charges	6
2	Introduction	8
3	Analyse - État de l'art	9
3.1	Protocoles existants	9
3.1.1	PGP	9
3.1.2	PEP	11
3.1.3	S/MIME	11
3.2	Implémentations existantes	11
3.2.1	Protonmail	11
3.2.2	Tutanota	12
3.3	Attaque existantes	13
3.3.1	Défauts webmail	13
3.3.2	EFAIL	13
3.4	Signal	13
3.4.1	Fonctionnement	13
3.4.2	Problèmes d'intégrations	13
3.5	Compromis	14
3.5.1	Résultats des recherches	14
3.6	Primitives	15
3.6.1	Primitives analysées	15
3.6.2	Primitive choisie	15
3.7	Recherches sur la primitive	15
3.7.1	Principes mathématiques	15
3.7.2	Schémas Certificateless de Chiffrement	15
3.7.3	Détails techniques	16
3.7.4	Schémas Certificateless de Signature	17
3.7.5	Détails techniques	17
4	Architecture / Design du protocole	18
4.1	Architecture globale	18
4.2	Acteurs	18
4.3	Fonctionnement Certificateless PKC	19

4.3.1	Chiffrement	19
4.3.2	Signature	19
4.4	Design du protocole	20
5	Implémentation	23
5.1	Choix d'implémentations	23
5.1.1	Langage	23
5.1.2	Librairie	23
5.1.3	Courbe utilisée	23
5.1.4	Dérivation de la clé AES	24
5.1.5	Fonctions de hachage - signature	24
5.1.6	Sérialisation des données	24
5.1.7	Enregistrement des clés publiques	24
5.2	Implémentation clés de chiffrement	24
5.2.1	Déroulement POC	25
6	Conclusion	30
	Bibliographie	31
A	Outils utilisés pour la compilation	34
A.1	RELIC Toolkit	34
A.2	Libsodium	34
B	Fichiers	35
B.1	Code du POC	35
B.2	Tableaux comparatifs	35

Chapitre 1

Cahier des charges

Résumé du problème

Les outils de chiffrement et de signature d'email actuels se résument principalement à S/MIME et à PGP.

Ces deux solutions sont anciennes, souffrent assez régulièrement de nouvelles vulnérabilités et ne proposent pas certaines propriétés cryptographiques qui pourraient être utiles (par exemple, la "forward secrecy". Le but de ce travail de bachelor est d'étudier quelles propriétés seraient utiles pour la sécurisation des emails, de proposer un nouveau protocole les implémentant et de développer un proof of concept.

Problématique

La problématique principale est résumée ci-dessus mais le principal problème c'est surtout que les technologies utilisées sont vieilles et elles souffrent de vulnérabilités par conception qui ont été mitigées en enlevant des options à l'utilisateur.

Solutions existantes

PGP, S/MIME, PEP, messagerie instantanée (Signal)...

Solutions possibles

Une solution possible est d'utiliser le système mis en place dans ce travail, cependant il faudrait une relecture et des analyses plus approfondies pour s'en assurer. Un des points possibles c'est de passer plus à de la messagerie instantanée dans la mesure du possible. Ou encore de s'orienter sur des nouvelles technologies comme PEP ou PGP mais en appliquant strictement les *Best Practices* et en se formant un peu sur leur utilisation qui n'est pas donnée à tout le monde.

Cahier des charges

Voici un résumé du cahier des charges sous formes d'objectifs à atteindre :

- Analyser les besoins d'un système d'E-mails actuel.
- Analyser et étudier les solutions de sécurité existantes.
- Comprendre et évaluer les propriétés cryptographiques défendues.
- Établir une liste des propriétés cryptographiques voulues pour un système de mails sécurisés.
- Trouver une primitive cryptographique satisfaisant les besoins énoncés et l'étudier pour en comprendre les bases et les besoins nécessaires en termes de sécurité.
- Établir la spécification pour un nouveau protocole en utilisant la primitive choisie.
- Faire un Proof Of Concept du protocole proposé.

Si le temps le permet :

- Comprendre plus en détails les mathématiques derrière la primitive utilisée.
- Faire un prototype de client mail utilisant une architecture mise en place pour le POC.

Déroulement

Tout d'abord je vais m'intéresser à faire une évaluation des concepts existants en messagerie sécurisée, tel que PGP et S/MIME pour les emails ou encore Signal pour la messagerie instantanée. Ayant vu ce qu'il se fait j'essaie de trouver une solution alternative pour le chiffrement et la signature d'emails. De là je vais conceptualiser un protocole et l'implémenter au sein d'un *Proof Of Concept*.

Livrables

Les livrables seront les suivants :

1. Une documentation contenant :
 - Une analyse de l'état de l'art
 - La décision qui découle de l'analyse
 - Spécifications
 - L'implémentation faites et les choix faits
 - Proof Of Concept
 - Les problèmes connus
2. Le code du *Proof Of Concept* fait, expliqué à l'aide de commentaires.

Chapitre 2

Introduction

Ce travail de Bachelor a pour but de sensibiliser à la vulnérabilité dans les systèmes actuels de messagerie électronique. Il propose aussi un nouveau protocole permettant de sécuriser ce type de messagerie à l'aide d'une primitive cryptographique peu implémentée, le *Certificateless Public Key Cryptography*. Ma démarche dans ce travail de bachelor est de voir si des solutions s'offrent à nous en considérons ce qui se fait sur le marché actuellement. En essayant d'améliorer les solutions actuelles proposées qui peuvent souffrir d'un manque de sécurité assez souvent ou (et plus souvent) un manque de simplicité d'utilisation.

Ce travail est découpé en plusieurs parties. En effet, on commence par une analyse de l'état de l'art, ce qui existe et voir pourquoi il faudrait de nouvelles solutions. Puis une présentation de la primitive cryptographique utilisée pour ma proposition dans ce travail. Enfin la présentation de l'architecture de mon protocole est une implémentation proposée en *Proof Of Concept* ainsi que les choix importants qui ont été faits en rapport à cette implémentation.

Chapitre 3

Analyse - État de l'art

La problématique principale est la difficulté d'utilisation des sécurités mises en places au-dessus des protocoles de base pour les emails. De plus des vulnérabilités (EFAIL en 3.3.2) qui réussissent à récupérer le texte clair a démontré que la sécurité n'était pas bien implémentée. Les vulnérabilités proviennent plus d'un défaut de conception inhérent aux mails. Je vais ici décrire les principaux problèmes trouvés sur PGP et S/MIME lors de mes tests d'utilisation et ce que j'ai trouvé durant mes recherches. De plus je vais analyser des systèmes de mails sécurisés tel que Protonmail et Tutanota.

3.1 Protocoles existants

Ici je vais regarder les protocoles existants afin de mettre en place des liaisons sécurisées de messagerie électronique.

3.1.1 PGP

Fonctionnement. PGP (Pretty Good Privacy ou Assez bonne confidentialité) est un moyen de chiffrer des données (mails, fichiers, ...) qui est beaucoup représenté lorsque l'on parle de sécurité email car c'est le plus utilisé avec S/MIME (c.f. 3.1.3). C'est une méthode de chiffrement hybride (utilise le chiffrement symétrique et asymétrique) qui fonctionne comme montré sur la Figure 3.1.

Ce fonctionnement hybride est défendu à cause de la lenteur et la non-praticité d'un chiffrement asymétrique sur un certain nombre de données. Ainsi en chiffrant uniquement la clé symétrique qui a servi à chiffrer le tout l'on peut déchiffrer bien plus rapidement et simplement le message (typiquement avec un chiffrement symétrique tel qu'AES qui a le droit à des instructions dédiées dans certains processeurs). Contrairement à des chiffrements asymétriques qui sont plus contraignants. Et l'on n'utilise pas directement le chiffrement symétrique car il a besoin d'un secret partagé dès le début de la communication. PGP utilise un système de clés... PGP est aussi critiqué pour son manque de "Forward Secrecy"...

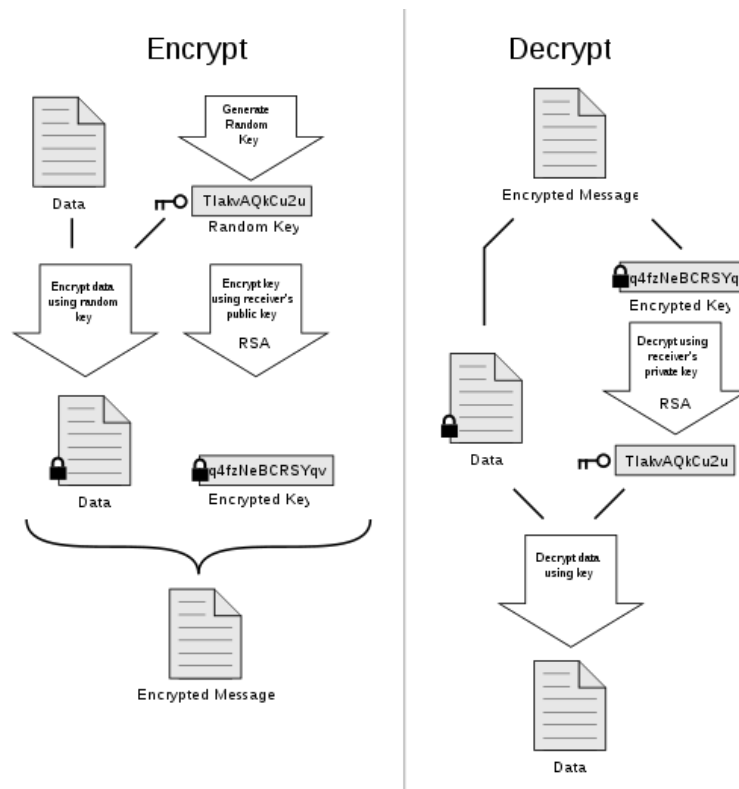


FIGURE 3.1 – Le fonctionnement global de PGP

Propriétés cryptographiques. Le problème qui est souvent reproché à PGP c'est qu'il n'implémentes pas de *Forward Secrecy*. La *Forward Secrecy* permet d'affirmer que si l'on a une brèche à un instant T , et qu'un attaquant récupère cette clé, il ne pourra pas déchiffrer les anciens messages. De plus, la gestion des clés PGP est très problématique, en effet lors de mes tests il était difficile de connecter un serveur de clés p. ex. Ou de recevoir une clé d'un correspondant pour la sauvegarder. Et même en la recevant, comment savoir si cette clé n'a pas été modifiée via un *MITM* p.ex.? -> utiliser un autre canal pour vérifier l'empreinte.

Web of Trust. Comment faire confiance a une clé -> surtout pour email -> Comment initialiser une confiance ?

Autocrypt. Autocrypt est une manière d'échanger des clés entres emails, ces échanges ne sont pas considérés sécurisé par la communauté (Wikipedia -> à creuser). C'est une façon de s'échanger des clés de manière automatisée mais pas forcément sécurisée (utilise les mails).

Utilisation. Pour mes tests j'ai fait en sorte de trouver l'utilisation la plus simple possible pour voir si un utilisateur lambda pouvait arriver à mettre en place ce genre de sécurité.

Il s'est avéré que cela était assez simple au départ, mais dès lors que l'on veut envoyer un mail chiffré à un correspondant cela se complique un peu. J'ai juste eu à installer un Add-On sur mon logiciel de messagerie (Thunderbird dans mon cas) qui s'appelle Enigmail. Ensuite Enigmail a généré mes clés PGP (de manière totalement opaque -> à creuser). Puis j'ai écrit un mail, en appuyant sur un petit cadenas mon mail partait chiffré et signé (uniquement si on a la clé du correspondant). Bien, cependant c'est très opaque et on ne sait pas ce qu'Enigmail et Autocrypt font réellement derrière les décors. L'utilisateur doit encore choisir s'il veut chiffrer ses mails ou non par contre il faut que le destinataire utilise PGP et que l'on ait sa clé publique. J'ai donc expérimenté à plus bas niveau ce qu'il se passait.

3.1.2 PEP

Citation. *By default, communications between pep peers always work end-to-end encrypted – no eavesdrop-ping in between shall be possible by design.*

Utilisation. pep assure un chiffrement de bout-en-bout par design, ils n'ont en effet pas de serveurs en soit et chiffrent à l'aide d'un *handshake* fait entre les deux personnes via des *trustwords*. Ce sont des mots qu'il faut vérifier entre les deux parties afin d'être sûr que la connexion est bien authentifiée. -> à creuser mais à priori PEP utilise PGP pour le chiffrement des messages.

3.1.3 S/MIME

Fonctionnement. Basé sur le même principe que PGP principalement, mais avec des certificats pour prouver la légitimité des clés publiques. Pour l'utilisation il faut se créer un certificat, plusieurs classes de confiance existent.

Propriétés cryptographiques.

Utilisation.

3.2 Implémentations existantes

3.2.1 Protonmail

Revendications. Protonmail revendique beaucoup de propriétés cryptographiques, tel que le zero-access encryption. Et l'end-to-end chiffrement + zero-knowledge pour les messages sécurisés, même avec leur fonctionnalité de (Chiffrement vers l'extérieur) utilisant AES256-GCM. Pour l'authentification Protonmail utilise une manière fortement sécurisée (SRP) pour ne pas avoir d'informations directes sur le mot de passe de l'utilisateur.

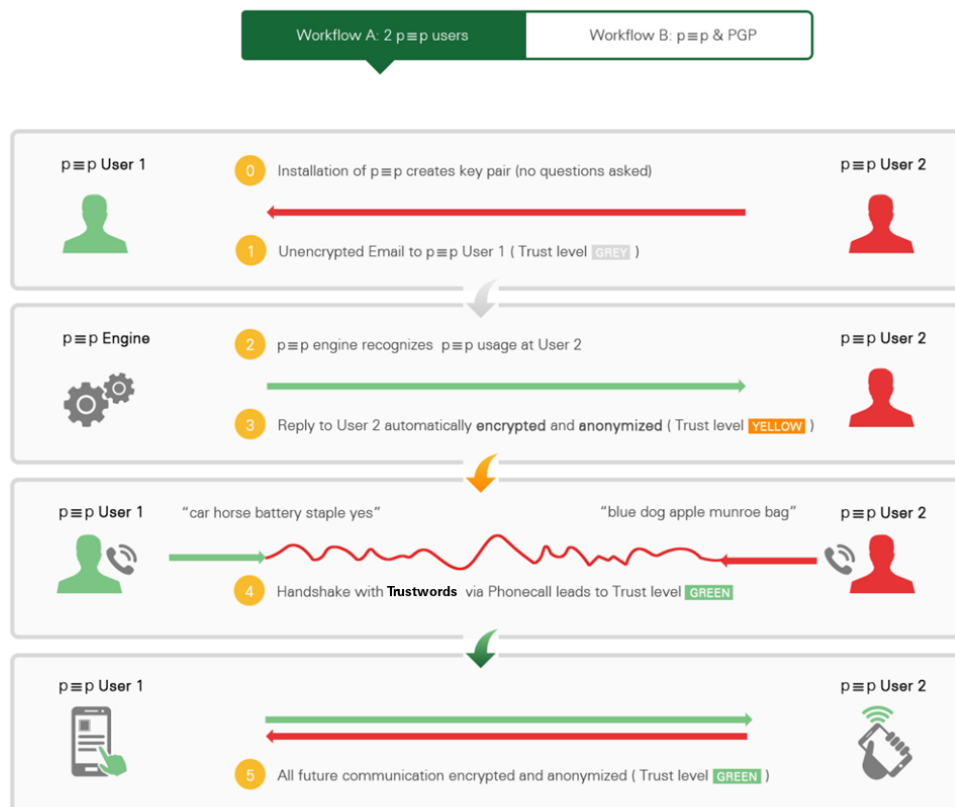


FIGURE 3.2 – Le fonctionnement global de pEp

Fonctionnement. Protonmail a plusieurs modes de fonctionnement dépendant du destinataire final. En effet de Protonmail à Protonmail les mails sont chiffrés à l'aide de PGP automatiquement. L'on peut utiliser Protonmail pour utiliser PGP si l'on a la clé de notre destinataire par exemple. Et l'on peut écrire un mail chiffré à quelqu'un qui n'utilise pas PGP grâce à une fonctionnalité de chiffrement vers l'extérieur. Cette fonctionnalité enverra une URL au destinataire qui, en la consultant, pourra déchiffrer le mail en utilisant un mot de passe communiqué de manière sécurisées entre les deux partis auparavant.

Open Source. Tout leur code est open-source afin d'avoir une validation externe, de plus ils ont un programme de Bug Bounty pour les chercheurs.

3.2.2 Tutanota

Fonctionnement. Tout ce que j'ai vu pour le moment c'est que Tutanota utilise AES128-CBC ? Mais dans PGP ou ailleurs ?

3.3 Attaque existantes

3.3.1 Défauts webmail

Selon un chercheur [6] l'infrastructure de Protonmail aurait des failles via son webmail. Mais son papier est en fait plus général et parle des webmails en règle général. Il part du principe que les serveurs de Protonmail ne sont pas des serveurs à faire confiance, pour ainsi prouver le zero-knowledge de Protonmail. Par contre, le fait qu'il ne peuvent pas être mis en confiance est un problème selon lui, car c'est ces serveurs qui vont délivrer le code d'OpenPGP afin de faire le chiffrement. Cela indique que si Protonmail était corrompu le fait d'avoir le code délivré par Protonmail pourrait avoir des effets néfastes. Comme p.ex l'extraction de la clé privée PGP. La conclusion est que dès le moment où vous avez utilisé une fois le webmail de protonmail la clé PGP est corrompue.

3.3.2 EFAIL

Malgré ces sécurités qui pourraient être mises en place à l'heure actuelle, une attaque nommée EFAIL [8] a été faite en 2018 et est toujours possible aujourd'hui (à vérifier / tester). En effet cette attaque a seulement été mitigée en évitant d'afficher les contenus HTML et les images dans boîtes mails de base. Car le problème vient de là principalement, des problèmes sont liés aussi aux modes de chiffrement utilisé (typiquement CBC et CFB) grâce à des "gadgets". Cette attaque permet en fait d'injecter une image dans l'HTML du message (typiquement dans les headers du mail), puis faire en sorte de récupérer le contenu du message déchiffrer dans un paramètre de l'URL.

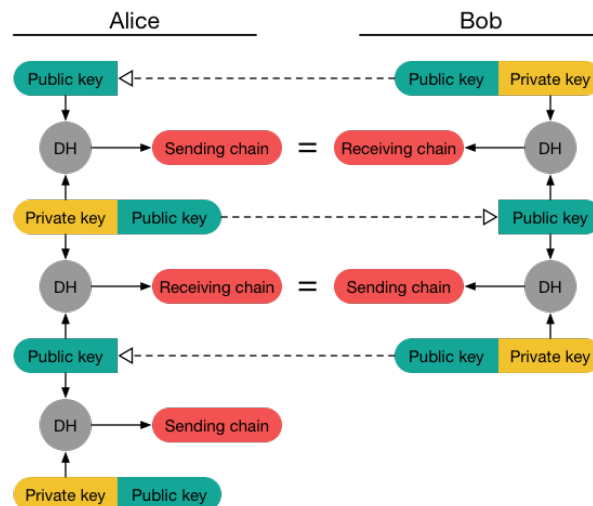
3.4 Signal

L'analyse s'est faite aussi pour la messagerie instantanée à cause de sa ressemblance avec la messagerie électronique.

3.4.1 Fonctionnement

3.4.2 Problèmes d'intégrations

Le problème avec le protocole Signal quant à mes besoins niveaux mails est la *forward secrecy* qui est très fort. En effet comme vu dans le chapitre précédent il utilise une clé par message grâce au *Double Ratchet*. Cependant ce fonctionnement comporte un gros problème en rapport aux mails, en effet si l'on veut pouvoir récupérer les anciens mails reçus/envoyés cela devient vraiment compliqué. En effet, la *forward secrecy* est une propriété utile dans un système de mail, mais faut pouvoir aussi récupérer les messages facilement si l'on connaît la clé privée.



3.5 Compromis

3.5.1 Résultats des recherches

3.6 Primitives

3.6.1 Primitives analysées

- Certificateless PKC [1]
- HIBE - Hierarchised Based encryption
- Identity based encryption

3.6.2 Primitive choisie

- Certificateless PKC [1]

Tout d'abord le HIBE aurait été compliqué à mettre en place aussi à cause de sa forte forward secrecy, la hiérarchie doit de plus être commune sur tout les appareils, ce qui prendrait probablement du temps à un nouvel arrivage d'appareil pour un compte donné. J'ai choisi cette primitive au final car elle proposait des propriétés très intéressantes pour ma manière d'implémenter dans les mails cela. En effet, similaire à de l'identity based avec un ID pour désigner une clé publique. Le problème avec l'identity based encryption c'est le fait que le serveur central génère la clé publique et la clé secrète de l'utilisateur, cela amène ce qu'on appelle le *key escrow* problème. C'est le fait qu'une entité connaisse à elle seule toutes les clés. Ce problème est résolu dans le certificateless en introduisant des *Partial Private Keys* permettant d'avoir une clé secrète partiellement générée par le serveur et par l'utilisateur.

3.7 Recherches sur la primitive

3.7.1 Principes mathématiques

Les variantes de *Certificateless Cryptography* choisies utilisent un concept appelé les *pairings* ou *bilinear map groups*. Des groupes tels que $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T$ d'un ordre premier p pour lesquels il existe un mapping $e : \mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_T$ avec les propriétés suivantes :

1. Bilinéarité : $e(g^a, h^b) = e(g, h)^{ab}$ pour tout $(g, h) \in \mathbb{G}_1 \times \mathbb{G}_2$ et $a, b \in \mathbb{Z}$;
2. Pas de dégénérescence : $e(g, h) \neq 1_{\mathbb{G}_T}$ tant que $g, h \neq 1_{\mathbb{G}_{1,2}}$;

3.7.2 Schémas Certificateless de Chiffrement

Pour choisir parmi les nombreux schémas existants en certificateless pour le chiffrement j'ai établi un tableau comparatif des différentes manières de faire, inspiré de [9]. En suivant ce tableau je me suis rendu compte que la construction de Dent-Libert-Paterson [4] était probablement la plus adaptée en vue des propriétés qu'elle présentait. Le tableau se trouve en annexe B.

3.7.3 Détails techniques

Les détails techniques sur le chiffrement avec la *Certificateless Cryptography*. Le chiffrement se base sur le problème difficile *The Decision 3-Party Diffie-Hellman Problem* (3-DDH).

C'est de décider si $T = g^{abc}$ ayant $(g^a, g^b, g^c, T) \in \mathbb{G}_4$.

Pour expliquer les détails techniques je vais ici montrer les calculs faits dans le schéma choisi [4] et les expliquer, cependant dans le schéma il est noté les calculs avec $\mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$ mais il est mentionné que c'est facilement adaptable pour $\mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_T$ ce que j'ai fait. De plus la conversion vers un groupe additif (travaillant sur les courbes elliptiques) est faite afin que les calculs ici puissent être lus avec mon implémentation :

Setup($1^k, n$) : Avec $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T$ avec un ordre $p > 2^k$. g est un générateur de \mathbb{G}_1 . Ensuite $g_1 = g * \gamma$ pour un $\gamma \leftarrow \mathbb{Z}_p^*$ aléatoire. Puis $g_2 \leftarrow \mathbb{G}_2$. Deux vecteurs (U,V) seront tirés aléatoirement dans \mathbb{G}_2^{n+1} en tant que fonctions de hash notés :

$$F_u(ID) = u' \sum_{i=1}^n u_j^{i_j} \quad \text{and} \quad F_v(w) = v' \sum_{i=1}^n v_j^{w_j}$$

L'on va aussi prendre une fonction de hash résistante aux collisions : $H : \{0, 1\}^* \rightarrow \{0, 1\}^n$.

Au final notre mpk (master public key) est :

$$mpk \leftarrow (g, g_1, g_2, U, V)$$

Et le msk (master seret key) est $msk \leftarrow g_2 * \gamma$.

Extract(mpk, γ, ID) : On prend $r \leftarrow \mathbb{Z}_p^*$ puis on retourne $d_{ID} \leftarrow (d_1, d_2) = (g_2 * \gamma + F_u(ID) * r, g * r)$

SetSec(mpk) : Retourne un secret aléatoirement choisi $x_{ID} \leftarrow \mathbb{Z}_p^*$.

SetPub(x_{ID}, mpk) : Retourne $pk_{ID} \leftarrow (X, Y) = (g * x_{ID}, g_1 * x_{ID})$.

SetPriv(x_{ID}, d_{ID}, mpk) : On choisit $r' \leftarrow \mathbb{Z}_p^*$ puis on reprends $(d_1, d_2) \leftarrow d_{ID}$ et l'on va prendre en secret key :

$$sk_{ID} \leftarrow (s_1, s_2) = (d_1 * x_{ID} + F_u(ID) * r', d_2 * x_{ID} + g * r')$$

Avec sk_{ID} étant la clé secrète de l'utilisateur, donnée par l'Extract (notre Partial Private Key) et la valeur secrète de SetSec.

Encrypt(m, pk_{ID}, ID, mpk) : Pour chiffrer $m \in \mathbb{G}_T$, l'on va reprendre $(X, Y) \leftarrow pk_{ID}$. Pour chiffrer ce message on va tiré aléatoirement $s \leftarrow \mathbb{Z}_p^*$ puis calculer :

$$C = (C_0, C_1, C_2, C_3) \leftarrow (m + e(Y, g_2) * s, g * s, F_u(ID) * s, F_v(w) * s)$$

Où $w \leftarrow H(C_0, C_1, C_2, ID, pk_{ID})$.

Decrypt(C, sk_{ID}, mpk) : L'on peut reprendre $(C_0, C_1, C_2, C_3) \leftarrow C$ et la clé privée $(s_1, s_2) \leftarrow sk_{ID}$. Afin d'accélérer le déchiffrement le calcul suivant peut être fait en tirant une valeur aléatoire $\alpha \leftarrow \mathbb{Z}_p^*$:

$$m = C_0 + \frac{e(s_2 + \alpha * g, C_2) * e(\alpha * g, C_3)}{e(C_1, s_1 + F_u(ID) * \alpha + F_v(w) * \alpha)}$$

Qui donnera m le texte en clair si le chiffré était bien formaté ou un élément aléatoire dans G_T .

3.7.4 Schémas Certificateless de Signature

Pour choisir parmi les nombreux schémas certificateless pour la signature j'ai établi un tableau comparatif des différentes manières de faire inspiré de [9]. En analysant les différentes possibilités dans ce tableau il y a peu de solutions se dégage, en effet l'on peut voir que beaucoup de schémas de signature sont cassés, mon choix s'est porté au final sur la construction de Zhang et Zhang [11] pour des signatures robustes en Certificateless. J'ai pris cette construction car elle est résistante au Malicious KGC (si le KGC a été setup avec des paramètres vulnérables) et en plus elle date de 2008 et n'a apparemment pas été cassée. Le tableau se trouve en annexe B.

3.7.5 Détails techniques

Les détails techniques sur la signature avec la *Certificateless Cryptography*. La signature se base sur le problème difficile *The Computational Diffie-Hellman Problem* (CDH).

Ayant P, aP, bP où a, b aléatoires $\in \mathbb{Z}_q^*$ il n'est pas possible de trouver abP .

Pour expliquer les détails techniques je vais ici montrer les calculs faits dans le schéma choisi [11] et les expliquer :

Setup :

Extract :

SetSec :

SetPub :

SetPriv :

Sign :

Verify :

Chapitre 4

Architecture / Design du protocole

Dans ce chapitre, je vais m'intéresser à expliquer le fonctionnement de la *certificateless cryptography* et démontrer comment je l'ai utilisée afin de l'intégrer à un protocole de chiffrement de mail.

4.1 Architecture globale

Dans cette Figure 4.1, je présente uniquement l'architecture globale pour bien représenter les différents acteurs présents dans le protocole et ainsi avoir une vue d'ensemble pour faciliter la compréhension.

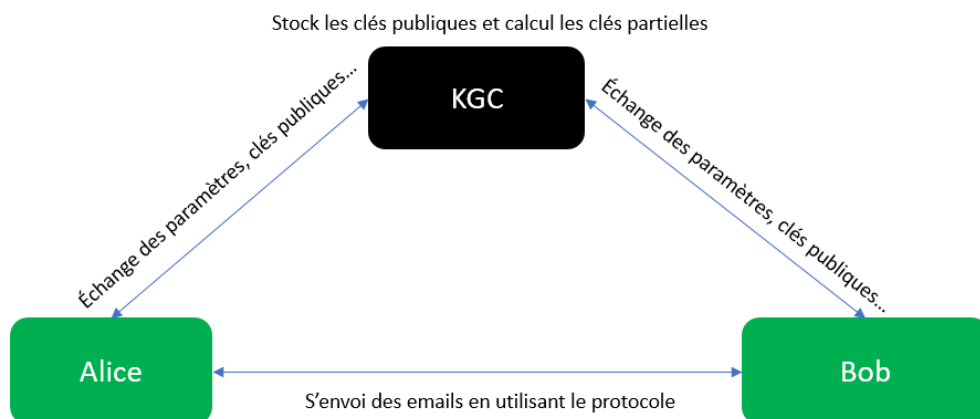


FIGURE 4.1 – Schéma global du protocole

4.2 Acteurs

Les parties impliquées sont les suivantes comme vu à la figure 4.1.

- Alice : L'envoyeur du message de mail en direction de Bob, ainsi elle doit discuter avec le KGC.
- Bob : Le destinataire du message, communique uniquement avec le KGC en ayant reçu le message d'Alice.
- KGC : Permet aux différents acteurs de pouvoir récupérer les Public Keys des clients, mais aussi recevoir les Partial Private Keys.

Ces partis sont les principaux présents dans un exemple de Certificateless Cryptography.

4.3 Fonctionnement Certificateless PKC

Je vais ici découper les différents algorithmes présent dans le certificateless public key cryptography. En passant par le chiffrement et la signature. Ces algorithmes seront accompagnés d'explications sur leur utilité. Les noms donnés aux algorithmes seront réutilisés ensuite pour les schémas afin de démontrer l'architecture du protocole mis en place. L'on peut voir des définitions spécifiques dans l'article sur lequel je me suis appuyé pour ce travail [4].

4.3.1 Chiffrement

Liste des différents algorithmes de *Certificateless Cryptography* et leur description.

- *Setup*. (seulement une fois par le KGC).
- *Partial-Private-Key-Extract*. Calcul d'une clé privée partielles lorsque qu'un client le demande pour identité donnée.
- *Set-Secret-Value*. Le client ne le fait qu'une fois pour tirer sa valeur secrète.
- *Set-Private-Key*. Le client combine ses clés partielles et sa clé secrète pour obtenir une clé privée afin de déchiffrer les message reçus, chiffrés avec une certaine identité.
- *Set-Public-Key*. Le client ne le fait qu'une fois, il calcule sa clé publique en fonction de sa valeur secrète.
- *Encrypt*. Chiffre un message avec la clé publique du destinataire et son identité.
- *Decrypt*. Déchiffres un message utilisant sa clé privée et l'identité utilisée pendant le chiffrement.

4.3.2 Signature

Pour la signature les algorithmes sont les mêmes avec une différence dans leur conception et évidemment le *Encrypt* et *Decrypt* sont remplacé par *Sign* et *Verify*. Dans la littérature certificateless les schémas de signatures sont beaucoup plus cassés que ceux de chiffrement apparemment (voir tableau). Il faut donc faire attention à suivre les schémas afin de vérifier que le schéma choisi ne soit pas mis à mal.

4.4 Design du protocole

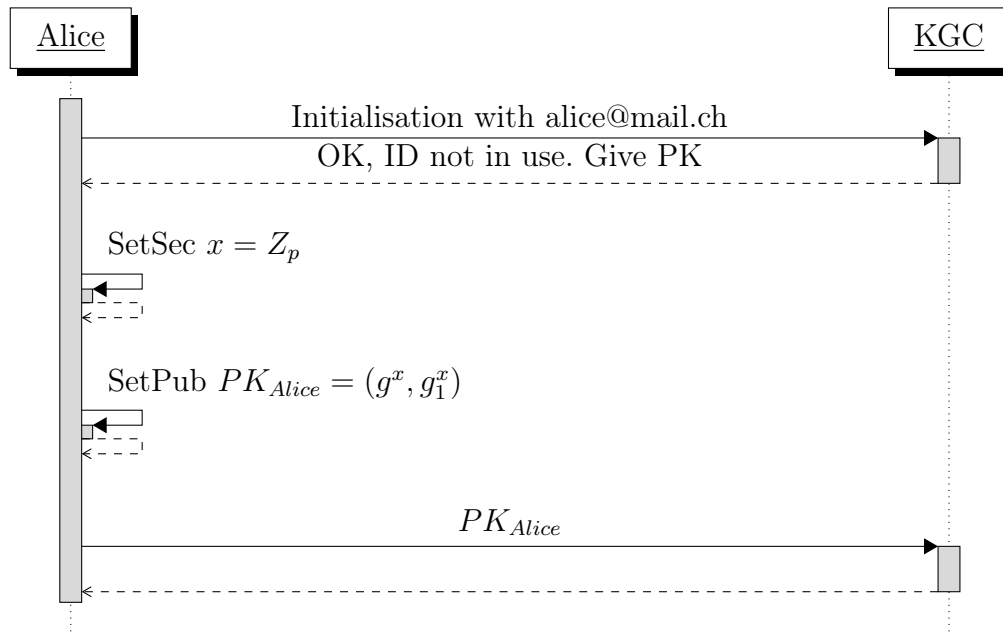


FIGURE 4.2 – Schéma de la première connexion

Dans la Figure 4.2 l'on voit la première connexion d'un utilisateur mais aussi le Setup du serveur, cette étape ne se fera qu'une fois dans la vie du KGC. Mais le Setup doit tout de même être fait pour la partie de signature et la partie de chiffrement. Ensuite Alice veut s'enregistrer auprès du KGC, ainsi le KGC lui renvoi les paramètres publics (mpkS et mpkE) si aucun utilisateur n'a déjà cet email. L'utilisateur va alors crée sa valeur secrète tirée aléatoire modulo p puis générer sa clé publique. Pour finir Alice envoie sa clé publique au KGC afin qu'il l'associe à son ID et puisse le donner aux personnes qui veulent envoyer un mail à Alice.

Dans la Figure 4.3 l'on voit comment se déroulerait l'envoi d'un message à Bob :

- Tout d'abord, Alice va récupérer le clé publique de Bob via son ID (aka email).
- Elle devra aussi récupérer sa clé privée partielle de signature pour créer ses clés privées afin de signer le message. Elle va le faire à l'aide de son ID et du même timestamp qu'utilisé pour la suite.
- Elle va ensuite tirer une valeur aléatoire dans G_t qui représentera sa clé AES pour la suite, elle va chiffrer cet élément à l'aide de la clé publique de Bob et de son ID complété par un timestamp. Ce timestamp sert à garder une certaine Forward Secrecy. Le cipher sera c' .
- Elle va calculer la signature du cipher donné (s' sur la figure)

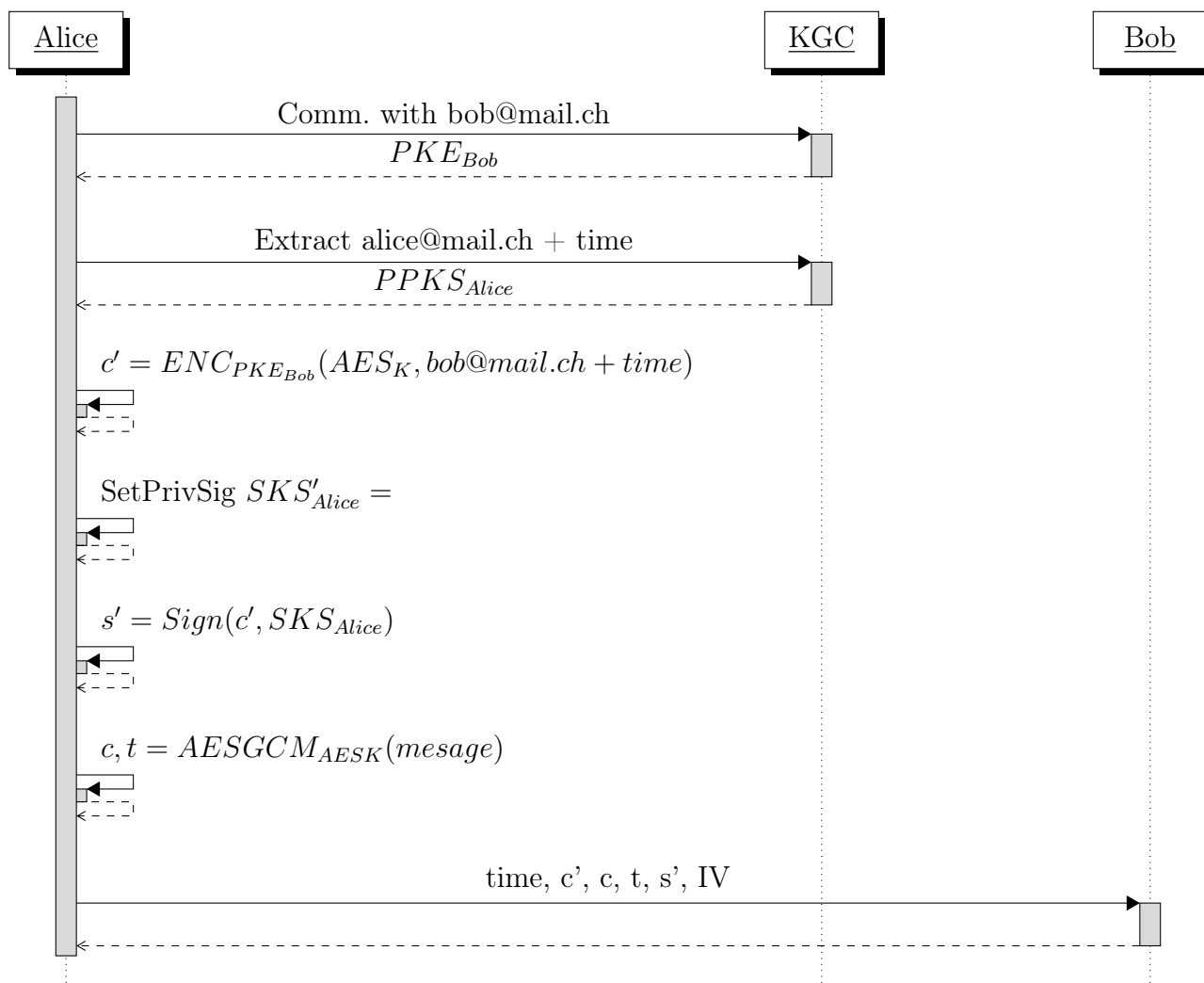


FIGURE 4.3 – Alice envoie un message à Bob

- Alice utilisera un chiffrement authentifié comme AES_GCM pour chiffrer et authentifier son mail à Bob, t pour le tag et c pour le cipher.
- Finalement elle va envoyer tout ces éléments à bob (à savoir, l'ID utilisé, c , c' , t , s' et l'IV utilisé pour AES_GCM).

Mais dans la Figure 4.4 l'on voit comment la réception du côté de Bob se déroulerait :

- A la réception la première chose à faire est de vérifier le cipher de la clé AES. Pour cela l'on va demander la clé publique d'Alice au KGC. Puis on va vérifier ce cipher c' à l'aide de sa signature s' .
- Ensuite Bob va récupérer sa clé privée partielle via le KGC en fournissant son ID avec le timestamp envoyé par Alice. Il va ainsi pouvoir former sa clé privée.
- Avec sa clé privée il va pouvoir déchiffrer c et obtenir la clé AES pour la suite.

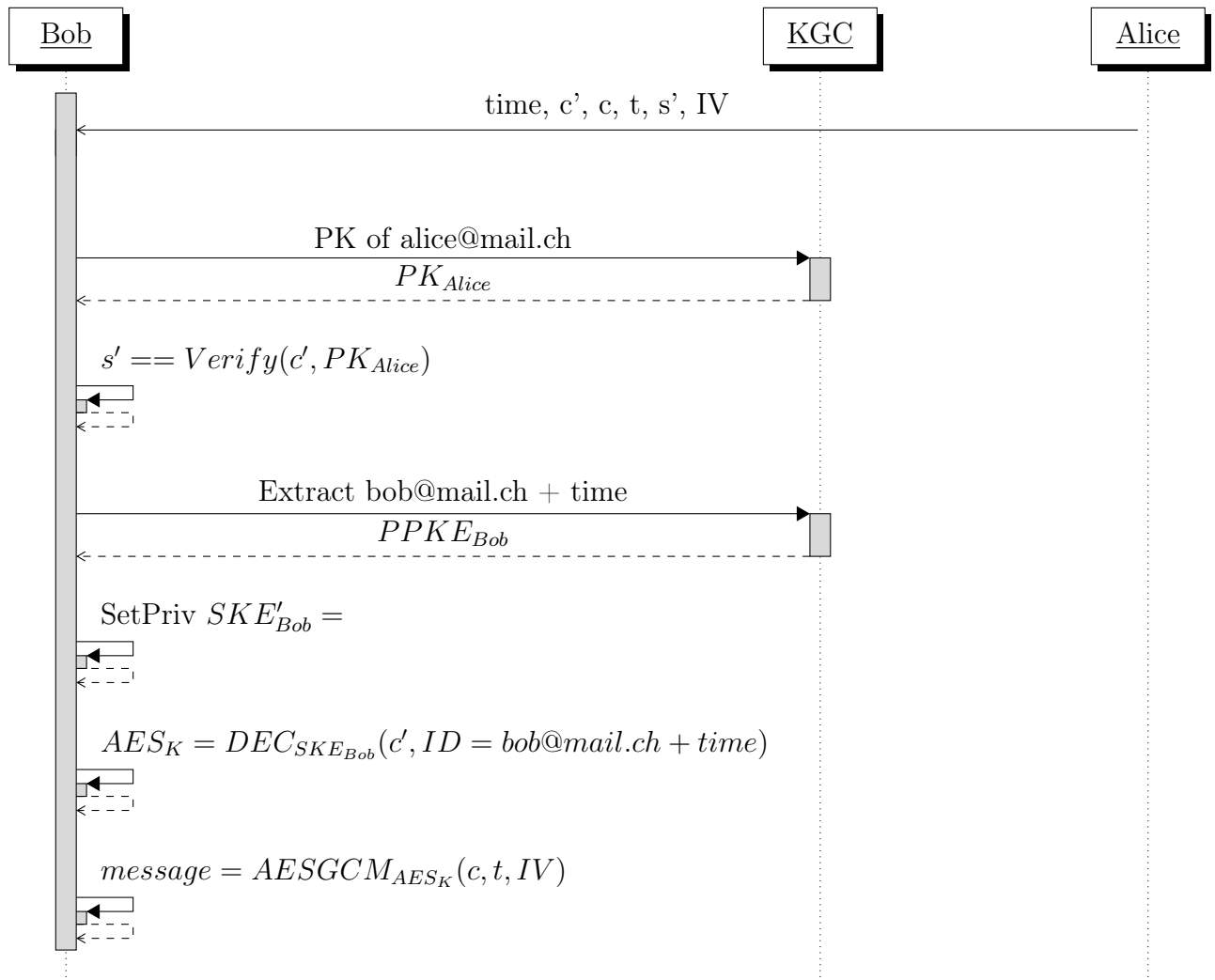


FIGURE 4.4 – Bob reçoit le message

- Une fois que l'on a la clé AES l'on peut simplement déchiffrer à l'aide d'AES_GCM c pour obtenir le message initial.

Chapitre 5

Implémentation

5.1 Choix d'implémentations

5.1.1 Langage

Au départ le choix du langage s'est porté sur sagemath (framework python) afin de mieux comprendre les différents calculs et faire un premier POC du chiffrement. Cependant l'implémentation du POC était lente et le changement d'algorithme pour les pairings était difficile. Je me suis donc orienté sur le C pour avoir de meilleures performances et pouvoir mieux gérer ma mémoire car c'est un point important dans des logiciels implémentant de la cryptographie. Pour pouvoir faire facilement des calculs sur les courbes elliptiques et les pairings en C il me fallait une librairie.

5.1.2 Librairie

La librairie utilisée est RELIC Toolkit [2], c'est une librairie en cours de développement qui se veut efficiente. Sa concurrence avec MIRACL m'a fait hésiter dans mon choix, mais MIRACL est plus codée en C++ avec des équivalences en C j'ai donc choisi RELIC.

5.1.3 Courbe utilisée

La courbe utilisée pour le POC est la BLS12-P381, en effet cette courbe est assez efficiente et compatible avec les pairings. De plus RELIC l'a dans ses options et fonctionne bien, elle a un niveau de sécurité de 128bits. Je voulais prendre une courbe avec une plus grande sécurité cependant RELIC ne l'a pas encore totalement implémenté (certains tests ne passent pas), mais la librairie étant toujours en cours de développement il faudrait suivre ça de près, le code ne changerait en effet pas.

5.1.4 Dérivation de la clé AES

Le but de mon schéma certificateless est de chiffrer et signer une clé AES qui permettra à mon message d'avoir un chiffrement authentifié. Pour cela il me faut dériver un élément de G_t en clé AES.

Pour cela j'ai utilisé une fonction permettant d'écrire sous forme compressée mon élément en bytes. Puis j'ai effectué un hachage avec SHA256 dessus, ainsi le résultat du hachage est une clé AES-256. La fonction de hachage doit être par conséquent cryptographiquement sûre.

5.1.5 Fonctions de hachage - signature

Pour le schéma de signature il nous faut plusieurs fonctions de hachage différentes (3 en fait). Pour appliquer cela j'ai utilisé la même méthode de mapping disponible dans RELIC pour mapper une char array (tableau de byte) à un point sur G_2 à savoir `g2_map`. Pour H_1 , la première fonction de hachage j'ai simplement utilisé cette fonction directement, mais pour H_2 et H_3 j'ai ajouté un byte devant les données à mapper respectivement les bytes '01' et '02'. Ceci afin de différencier les hash générés. Ceci s'appelle du 'Domain Hash Separation'.

5.1.6 Sérialisation des données

Pour la sérialisation des données, typiquement les clés publiques et les clés privées partielles envoyées en réseau ou les clés publiques enregistrées dans les fichiers par exemple, j'ai utilisé la librairie `binn`¹. Cela permet de packer facilement des données binaires, pour cela RELIC met à disposition des méthodes `g1_write_bin` `g1_read_bin` qui a permis de faire ces enregistrements binaires. Ainsi les transferts de données sont simplifiés.

5.1.7 Enregistrement des clés publiques

Pour les enregistrements et récupération des clés publiques pour le KGC j'ai simplement mis en place une structure de fichiers simples : un dossier signature et un encryption. De là je crée un fichier par utilisateur avec leur ID comme nom de fichier, et met la clé publique dedans, ainsi lorsque l'on doit récupérer des clés publiques c'est assez simple à retrouver les clés.

5.2 Implémentation clés de chiffrement

Pour pouvoir implémenter ce schéma de chiffrement et signature certificateless dans un système hybride il a fallu penser à une manière d'encapsuler la clé et les données. Pour cela j'ai essayé de faire un système comparable à la figure 5.1.

1. <https://github.com/liteserver/binn>

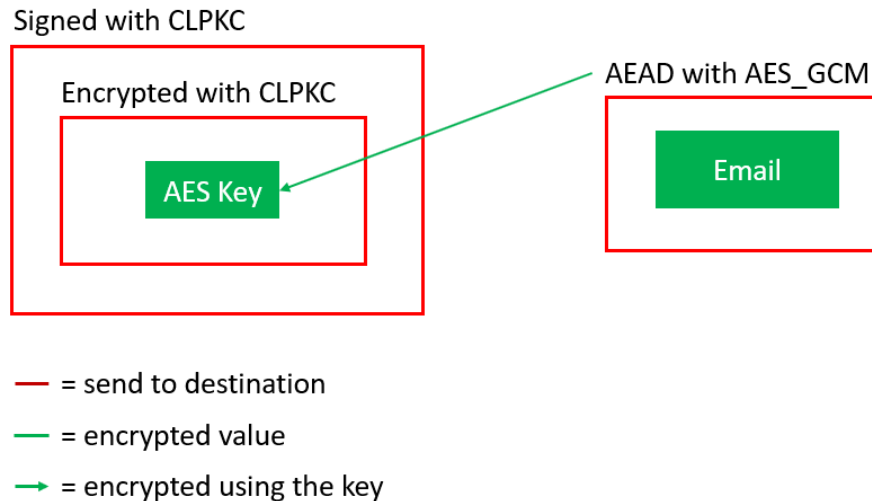


FIGURE 5.1 – Schéma encapsulation des données

5.2.1 Déroulement POC

Je mets ici temporairement le code du main afin de voir le déroulement global d'un échange et voir ce qui a été implémenté jusque là.

```

1  #include "cipherPOC.h"
2  #include "signaturePOC.h"
3  #include "sodium.h"
4
5  // Utils function for encrypting / decrypting AES_GCM
6  char* encrypt_message(const char* m, unsigned char* key, unsigned char* nonce, unsigned
    ↪ char* cipher, unsigned long long* cipher_len, const size_t* m_len){
7      randombytes_buf(nonce, sizeof nonce);
8      crypto_aead_aes256gcm_encrypt(cipher, cipher_len, m, *m_len, NULL, 0, NULL, nonce,
    ↪ key);
9  }
10
11 void decrypt_message(unsigned char* decrypted, unsigned char* cipher, unsigned char*
    ↪ nonce, unsigned char* key, unsigned long long cipher_len){
12     unsigned long long decrypted_len;
13     if (cipher_len < crypto_aead_aes256gcm_BYTES ||
14         crypto_aead_aes256gcm_decrypt(decrypted, &decrypted_len,
15                                     NULL,
16                                     cipher, cipher_len,
17                                     NULL,
18                                     0,
19                                     nonce, key) != 0) {
20         /* message forged! */
21         printf("Message not correctly authenticated ! Aborting decryption...\n");
22     }
23 }
  
```

```
24
25 void get_key(char *aesk, gt_t originalM) {
26     int sizeAESK = gt_size_bin(originalM,1);
27     char aeskBin [sizeAESK];
28     gt_write_bin(aeskBin, sizeAESK, originalM, 1);
29     md_map_sh256(aesk, aeskBin, sizeAESK);
30     printf("AES Key : ");
31     for(int i=0;i < 32;i++)
32         printf("%02X",(unsigned char)aesk[i]);
33     printf("\n");
34 }
35
36 int main() {
37     if(core_init() == RLC_ERR){
38         printf("RELIC INIT ERROR !\n");
39     }
40     if(sodium_init() < 0) {
41         printf("LIBSODIUM INIT ERROR !\n");
42     }
43     if(pc_param_set_any() == RLC_OK){
44         // Server doing this once
45         pc_param_print();
46         // Setup the encrypting and signing parameters for KGC
47
48         printf("Security : %d\n", pc_param_level());
49
50         // MPK struct, Master Public Key structure to store
51         mpkStruct mpkSession;
52         mpkStructSig mpkSignature;
53
54         // Master secret key of KGC for encrypting
55         g2_t msk;
56         g2_null(msk)
57         g2_new(msk)
58
59         setup(256, &mpkSession, &msk);
60
61         // Master key of KGC for signing
62         bn_t masterSecret;
63         bn_null(masterSecret)
64         bn_new(masterSecret)
65
66         setupSig(256, &mpkSignature, &masterSecret);
67         printf("Setup successful !\n");
68         // -----
69         // At this point, setup of KGC for encrypting and signing is successful
70
71         // Now we can go for user's private keys (encrypting and signing)
72         bn_t x;
73         setSec(&x);
74
```

```
75     bn_t xSig;
76     setSecSig(&xSig);
77     // -----
78     // Private keys set
79
80     // Now we can go to set Public keys for both signing and encrypting
81     PK myPK;
82     setPub(x, mpkSession, &myPK);
83
84     PKSig myPKSig;
85     setPubSig(xSig, mpkSignature, &myPKSig);
86     // -----
87     // Public keys set
88
89
90     // The other user takes ID of the destination and PK to encrypt his message
91     // With the final version we will need to append a timestamp on the ID
92     char ID[] = "mickael.bonjour@hotmail.fr";
93
94     gt_t AESK;
95     gt_null(AESK);
96     gt_new(AESK);
97     // For now we take m (AES Key) randomly from Gt
98     gt_rand(AESK);
99
100     unsigned char aesk [crypto_secretbox_KEYBYTES];
101     get_key(aesk, AESK);
102
103     char* m = "This message will be encrypted";
104     printf("Message : %s\n", m);
105     unsigned char nonceAES[crypto_aead_aes256gcm_NPUBBYTES];
106     size_t m_len = strlen(m);
107     unsigned long long cipher_len;
108     unsigned char ciphertextAES[m_len + crypto_aead_aes256gcm_ABYTES];
109     encrypt_message(m, aesk, nonceAES, ciphertextAES, &cipher_len, &m_len);
110     printf("Encrypted message : %s\n", ciphertextAES);
111
112     cipher c;
113     encrypt(AESK, myPK, ID, mpkSession, &c);
114
115     // For the signature we need our PPK
116     PPSig myPartialKeysSig;
117
118     //The sender needs to extract (via KGC) and setPriv to get his private key and
119     ↪ sign the message
120     extractSig(mpkSignature, masterSecret, ID , &myPartialKeysSig);
121
122     // Computes Secret User Keys for Signature
123     SKSig mySecretKeysSig;
124     setPrivSig(xSig, myPartialKeysSig, mpkSignature, ID, &mySecretKeysSig);
```

```
125 // Computes the message to sign, so the cipher struct
126 int c0Size = gt_size_bin(c.c0,1);
127 int c1Size = g1_size_bin(c.c1, 1);
128 int c2Size = g2_size_bin(c.c2, 1);
129 int c3Size = g2_size_bin(c.c3, 1);
130 uint8_t mSig[c0Size+c1Size+c2Size+c3Size];
131 gt_write_bin(mSig, c0Size, c.c0, 1);
132 g1_write_bin(&mSig[c0Size], c1Size, c.c1, 1);
133 g2_write_bin(&mSig[c0Size + c1Size], c2Size, c.c2, 1);
134 g2_write_bin(&mSig[c0Size + c1Size + c2Size], c3Size, c.c3, 1);
135
136 // Structure of an signature
137 signature s;
138 // We can sign using our private keys and public ones
139 sign(mSig, mySecretKeysSig, myPKSig, ID, mpkSignature, &s);
140 // -----
141 // Now the message is encrypted and authenticated with an AES Key and the key is
142 //   ↪ encrypted and signed using CLPKC
143 // -----
144
145
146 // We can go for decrypting and verification
147 // For this we need our Partial Private Keys with the ID used to encrypt the
148 //   ↪ message
149
150 // We can verify directly with the public keys of the sender
151 int test = verify(s, myPKSig, mpkSignature, ID, mSig);
152 printf("\nVerification of the key (0 if correct 1 if not) : %d\n", test);
153 // if the verif is ok we can continue, otherwise we can stop here
154 if(test == 0) {
155     PPK myPartialKeys;
156     g2_null(myPartialKeys->d1)
157     g2_new(myPartialKeys->d1)
158
159     g1_null(myPartialKeys->d2)
160     g1_new(myPartialKeys->d2)
161
162     //The receiver needs to extract (via KGC) and setPriv to get his private
163     //   ↪ key and decrypt the cipher
164     extract(mpkSession, msk, ID, &myPartialKeys);
165
166     // Computes Secret User Keys
167     SK mySecretKeys;
168     g2_null(mySecretKeys->s1)
169     g2_new(mySecretKeys->s1)
170
171     g1_null(mySecretKeys->s2)
172     g1_new(mySecretKeys->s2)
173     setPriv(x, myPartialKeys, mpkSession, ID, &mySecretKeys);
```

```
173      // We can decrypt now
174      gt_t decryptedMessage;
175      gt_null(decryptedMessage)
176      gt_new(decryptedMessage)
177      decrypt(c, mySecretKeys, myPK, mpkSession, ID, &decryptedMessage);
178
179      char aeskDecrypted[crypto_secretbox_KEYBYTES];
180      get_key(aeskDecrypted, decryptedMessage);
181
182      unsigned char decrypted[m_len];
183      decrypt_message(decrypted, ciphertextAES, nonceAES, aeskDecrypted,
184                     ⇐ cipher_len);
185      printf("Decrypted message : %s\n", decrypted);
186  }
187
188      // For test purposes
189      // We change the message to see the signature not being correct again
190      unsigned char* mSigCorrupt = "The message to be signed !!";
191      printf("Message changed to simulate corruption\n");
192
193      // We can verify now with the public keys of the sender
194      test = verify(s, myPKSig, mpkSignature, ID, mSigCorrupt);
195      printf("Verification (0 if correct 1 if not) : %d\n", test);
196  }
197  core_clean();
```

Chapitre 6

Conclusion

Bibliographie

- [1] Sattam S. Al-Riyami and Kenneth G. Paterson. Certificateless public key cryptography. In Chi-Sung Lai, editor, *Advances in Cryptology - ASIACRYPT 2003, 9th International Conference on the Theory and Application of Cryptology and Information Security, Taipei, Taiwan, November 30 - December 4, 2003, Proceedings*, volume 2894 of *Lecture Notes in Computer Science*, pages 452–473. Springer, 2003.
- [2] D. F. Aranha, C. P. L. Gouvêa, T. Markmann, R. S. Wahby, and K. Liao. RELIC is an Efficient Library for Cryptography. <https://github.com/relic-toolkit/relic>.
- [3] Suresh Kumar Balakrishnan and V. P. Jagathy Raj. Practical implementation of a secure email system using certificateless cryptography and domain name system. *I. J. Network Security*, 18(1) :99–107, 2016.
- [4] Alexander W. Dent, Benoît Libert, and Kenneth G. Paterson. Certificateless encryption schemes strongly secure in the standard model. In Ronald Cramer, editor, *Public Key Cryptography - PKC 2008, 11th International Workshop on Practice and Theory in Public-Key Cryptography, Barcelona, Spain, March 9-12, 2008. Proceedings*, volume 4939 of *Lecture Notes in Computer Science*, pages 344–359. Springer, 2008.
- [5] Yee-Lee Er, Wei-Chuen Yau, Syh-Yuan Tan, and Bok-Min Goi. Email encryption system using certificateless public key encryption scheme. In James Jong Hyuk Park, Jongsung Kim, Deqing Zou, and Yang Sun Lee, editors, *Information Technology Convergence, Secure and Trust Computing, and Data Management - ITCS 2012 & STA 2012, Gwangju, Korea, September 6-8, 2012*, volume 180 of *Lecture Notes in Electrical Engineering*, pages 179–186. Springer, 2012.
- [6] Nadim Kobeissi. An analysis of the protonmail cryptographic architecture. *IACR Cryptol. ePrint Arch.*, 2018 :1121, 2018.
- [7] Trevor Perrin and Moxie Marlinspike. The Double Ratchet Algorithm. Technical report, 2016.
- [8] Damian Poddebniak, Christian Dresen, Jens Müller, Fabian Ising, Sebastian Schinzel, Simon Friedberger, Juraj Somorovsky, and Jörg Schwenk. Efail : Breaking S/MIME and openpgp email encryption using exfiltration channels. In William Enck and Adrienne Porter Felt, editors, *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018*, pages 549–566. USENIX Association, 2018.

- [9] Hu Xiong, Zhen Qin, and Athanasios V. Vasilakos. *Introduction to Certificateless Cryptography*. CRC Press, Inc., USA, 2016.
- [10] Jyh-haw Yeh, Srisarguru Sridhar, Gaby G. Dagher, Hung-Min Sun, Ning Shen, and Kathleen Dakota White. A certificateless one-way group key agreement protocol for end-to-end email encryption. In *23rd IEEE Pacific Rim International Symposium on Dependable Computing, PRDC 2018, Taipei, Taiwan, December 4-7, 2018*, pages 34–43. IEEE, 2018.
- [11] Lei Zhang and Futai Zhang. A new provably secure certificateless signature scheme. In *Proceedings of IEEE International Conference on Communications, ICC 2008, Beijing, China, 19-23 May 2008*, pages 1685–1689. IEEE, 2008.

Table des figures

3.1	Le fonctionnement global de PGP	10
3.2	Le fonctionnement global de pEp	12
3.3	Schéma fonctionnement de Signal[7]	14
4.1	Schéma global du protocole	18
4.2	Schéma de la première connexion	20
4.3	Alice envoie un message à Bob	21
4.4	Bob reçoit le message	22
5.1	Schéma encapsulation des données	25

Annexe A

Outils utilisés pour la compilation

A.1 RELIC Toolkit

Pour pouvoir faire des calculs de *Pairings* et sur des courbes elliptiques je me suis fier à RELIC Toolkit [2] qui est une librairie C permettant ce genre de calculs assez simplement. Cette librairie demande à être compilée avec une certaine courbe et certaines options (typiquement fonction de hachage et autres...). Des presets existent et c'est donc ce que j'ai utilisé pour ce POC. Cela demande donc de fournir la librairie précompilée avec les bonnes options pour l'utilisateur. L'inconvénient c'est donc que pour mettre à jour une courbe il va falloir recompiler toute la librairie et la fournir à l'utilisateur, néanmoins on n'aura pas à changer de code.

A.2 Libsodium

Pour faire du chiffrement authentifié j'ai utilisé libsodium¹, en effet, m'étant un peu familiarisé avec la librairie il m'a semblé être le choix le plus évident en plus de fournir des méthodes de chiffrement simples à mettre en place. Nécessite d'avoir libsodium en librairie liée.

1. <https://libsodium.gitbook.io/doc/>

Annexe B

Fichiers

Je liste ici les fichiers annexes à mon rapport, ce qu'ils contiennent et comment les utiliser si besoin.

B.1 Code du POC

Le code est en annexe du rapport avec le nom *POCCertificatelessCryptography*. Le *README.md* présent à la source devrait être suffisant pour compiler soit même le code. Le code est très commenté au niveau du *main.c* pour bien montrer les différentes étapes telles qu'elle pourraient arriver dans une implémentation finale.

B.2 Tableaux comparatifs

Les tableaux comparatifs cité dans le chapitre 3 apparaissent sous forme de feuille dans un fichier excel se nommant *ComparatifsCLPKCSchemes.xlsx*. La feuille nommée CLEs contient un comparatif des schémas de chiffrement tandis que la feuille CLSs contient un comparatif des schémas de signatures.