



**ETUDE DE LA CONFIDENTIALITÉ
D'UN LOGICIEL DE TRAITEMENT DE DONNEES
EN ENTREPRISE**

Juin 2015

à INTRACHERCHE



Par Abdourahamane Conde
(Élève Ingénieur ISAE promotion 69)
Tuteur ISAE : Frédéric Lachaud
Tuteur Entreprise : Rémi Tournier

Table des matières

<u>INTRODUCTION.....</u>	<u>2</u>
<u>CADRE D'ETUDE.....</u>	<u>3</u>
<u>PHASE DE TEST.....</u>	<u>3</u>
PRESENTATION DE LA PROCEDURE.....	3
DEROULEMENT DE LA PROCEDURE.....	3
1 – Test pendant l'installation.....	4
2 – Test pendant l'indexation.....	4
3 – Test pendant la recherche.....	4
<u>PRESENTATION DES RESULTATS.....</u>	<u>4</u>
<u>ANALYSE DES RESULTATS.....</u>	<u>5</u>
<u>CONCLUSION.....</u>	<u>5</u>

INTRODUCTION

Dans le cadre de notre étude, nous nous sommes penchés sur la confidentialité des données en entreprise notamment sur les risques liés à l'utilisation de programmes de gestion et de traitement de données.

Avec la dématérialisation de l'information les grandes entreprises amassent de plus en plus de données numériques. Le traitement de ces données étant fastidieux elles utilisent pour ce faire en général des applications spécialisées. Ces applications réalisent une indexation de l'ensemble des documents de l'entreprise ce qui permet de les retrouver de manière plus simple et plus rapide.

L'utilisation de ce type de programme offre un gain de temps énorme lors de la recherche d'information mais peut soulever cependant quelques interrogations. Notamment celle de la confidentialité des données. En effet, les programmes en question traitent des documents contenant des informations sensibles et qui doivent rester confidentielles, il faudrait donc s'assurer que celles-ci ne fuient pas vers l'extérieur par le biais de l'application.

Nous allons réaliser un test sur l'un des quelques programmes de ce genre présents sur le marché afin d'analyser l'ensemble des données envoyées et reçues au cours de son fonctionnement. Le programme choisi pour notre étude est celui développé par la société IntraCherche basée à Toulouse. Choix axé sur l'accessibilité de la société par l'auteur de cette étude, lui-même étudiant de cette ville. Nous nous sommes ainsi procurés le logiciel en question auprès de son concepteur en lui signifiant le but de notre étude.

L'objectif est donc de réaliser une série de test sur ce logiciel et ensuite d'analyser les résultats afin de pouvoir en tirer des conclusions intéressantes. Avant tout fixons le cadre de notre étude.

CADRE D'ETUDE

Commençons par une petite explication de ce qu'est IntraCherche et comment il fonctionne.

IntraCherche est un logiciel de traitement avancé de données. Son mode de fonctionnement est le suivant : il est constitué de deux modules, un indexeur et plusieurs fouilleurs.

L'indexeur est celui qui réalise l'indexation. Il est centralisé, on peut l'installer sur le serveur de répertoires partagés de l'entreprise ou même sur l'ordinateur d'un employé quelconque. Ce module va donc passer en revue tous les documents des répertoires partagés et réaliser un index rassemblant tous les mots-clés. Cette phase est la plus longue (comparée à la recherche avec le fouilleur) mais elle demeure essentielle.

Ensuite grâce aux fouilleurs installés sur les différents ordinateurs il sera possible de retrouver n'importe quel document juste en entrant des mots ou des expressions (comme on le fait sur Google). Le fouilleur va ainsi aller chercher dans l'index tous les documents qui ont un rapport avec les mots ou expression en question.

Toute cette procédure se déroule dans le réseau interne de l'entreprise et aucun document n'est censé se retrouver à l'extérieur de ce réseau. C'est en cela qu'une étude de confidentialité s'impose.

Comment donc allons nous procéder pour étudier sa confidentialité ?

C'est simple ce que l'on va faire c'est installer IntraCherche sur un ordinateur de test. On va y mettre l'index et le fouilleur afin de simuler un réseau local puis nous allons utiliser un logiciel traceur (sniffer en Anglais), qui va scanner toutes les requêtes en entrée et en sortie de notre ordinateur pendant toutes les phases de fonctionnement du logiciel. Dans l'idéal on ne devrait en intercepter aucune mais si cela arrive nous l'analyserons afin de savoir ce qu'elle contient.

Maintenant que le cadre est fixé, nous pouvons désormais débiter la phase de Test

PHASE DE TEST

PRESENTATION DE LA PROCEDURE

Comme mentionné plus haut ce test consiste à utiliser un logiciel traceur afin de scanner le trafic pendant toute la phase d'installation et d'utilisation du programme IntraCherche.

Les données du test sont les suivantes:

- **durée de la procédure:** 1 semaine – 2h/jour
- **environnement:** Windows 8
- **logiciel de traçage:** Microsoft Network Monitor
- **but du test:** repérer un trafic et analyser la nature des données envoyées
- **types de documents indexés :** PDF, PDF scannés, Textes, Présentations.

DEROULEMENT DE LA PROCEDURE

La procédure de test s'est faite en plusieurs étapes:

- Test au cours de l'installation
- Test pendant l'indexation
- Test pendant la recherche

1 – Test pendant l'installation

Pour ce test nous avons commencé par activer notre logiciel traceur. Ensuite nous avons lancé l'installation du logiciel IntraCherche. Notre traceur a donc scanné toutes les entrées et sorties pendant la période d'installation. L'opération a duré le temps de l'installation (2 à 3 minutes).

2 – Test pendant l'indexation

Avec la même méthode nous avons scanné les entrées et sorties pendant la période d'indexation. L'opération a duré le temps de l'indexation (10 à 15 minutes).

3 – Test pendant la recherche

Nous avons fait de même pendant la phase de recherche tout en testant plusieurs cas d'utilisation du logiciel; Lancement d'une requête, ouverture de fichiers, ouverture de l'aide, clics sur les liens externes, fermeture de l'application. L'opération a duré deux heures.

L'ensemble des résultats de ces tests sont répertoriés dans le tableau suivant.

PRESENTATION DES RESULTATS

PHASE	Req. envoyée	Req. reçue	PAR	VERS	Type de requête
INSTALLATION	1	0	intracherche	http://intracherche.com	Message de bienvenue
INDEXATION	1	0	intracherche	http://intracherche.com	Lien de redirection
UTILISATION	1	0	intracherche	http://intracherche.com	Lien de redirection

ANALYSE DES RESULTATS

Les résultats nous montrent qu'il y a trois requêtes envoyées au cours de ces trois phases vers le site web intracherche.com situé sur un serveur externe au domaine d'étude.

Nous avons donc analysé le contenu de ces requêtes:

- La première requête envoyée au cours de l'installation contenait des informations en vue d'une personnalisation du message de fin d'installation.
- Les deux autres envoyées au cours de l'utilisation par le fouilleur et l'indexeur ne contenaient également aucune donnée sensible. Ce n'était en effet que des liens de redirection.

CONCLUSION

A l'issue de la série de test nous n'avons pas intercepté de données sensibles envoyées à l'extérieur. Il apparaît donc au terme de notre étude que le logiciel pris en exemple assure bien le respect de la confidentialité des documents traités.