Aide-mémoire pour le téléchargement de fichiers

Introduction

Le téléchargement de fichiers devient une fonctionnalité de plus en plus essentielle dans toute application, où l'utilisateur peut télécharger ses photos, son CV ou une vidéo présentant un projet sur lequel il travaille. L'application doit être capable de repousser les fichiers frauduleux et malveillants afin d'assurer la sécurité de l'application et des utilisateurs.

En résumé, les principes suivants doivent être respectés pour garantir la sécurité du téléchargement de fichiers :

- **Répertorier les extensions autorisées. N'autoriser que les extensions sûres et essentielles au fonctionnement de l'application.**
  - **S'assurer que la validation des entrées est appliquée avant de valider les extensions.**
- **Validez le type de fichier, ne vous fiez pas à l'en-tête Content-Type car il peut être falsifié**
- **Modifiez le nom du fichier pour lui donner un nom généré par l'application.**
- **Définissez une limite de longueur pour le nom de fichier. Si possible, limitez les caractères autorisés.**
- **Définissez une limite de taille de fichier.**
- **Autorisez uniquement les utilisateurs autorisés à télécharger des fichiers**
- **Stockez les fichiers sur un autre serveur. Si cela n'est pas possible, stockez-les en dehors du répertoire web.**
  - **Dans le cas d'un accès public aux fichiers, utilisez un gestionnaire qui est mappé aux noms de fichiers dans l'application (someid -> file.ext)**
- **Exécutez le fichier à l'aide d'un antivirus ou d'un bac à sable, si disponible, afin de vérifier qu'il ne contient pas de données malveillantes.**
- **Exécutez le fichier via CDR (Content Disarm & Reconstruct) si le type le permet (PDF, DOCX, etc.)**
- **Assurez-vous que toutes les bibliothèques utilisées sont configurées de manière sécurisée et maintenues à jour.**
- **Protégez le téléchargement de fichiers contre les attaques CSRF**

Menaces liées au téléchargement de fichiers

Afin d'évaluer et de savoir exactement quels contrôles mettre en œuvre, il est essentiel de connaître les risques auxquels vous êtes confronté pour protéger vos actifs. Les sections suivantes présentent les risques liés à la fonctionnalité de téléchargement de fichiers.

Fichiers malveillants

L'attaquant transmet un fichier à des fins malveillantes, par exemple :

1. Exploiter les vulnérabilités du parseur de fichiers ou du module de traitement (*par exemple,* ImageTrick Exploit, XXE)
2. Utiliser le fichier à des fins de phishing (*par exemple,* formulaire de candidature)
3. Envoyer des bombes ZIP, des bombes XML (également appelées « billion laughs attack ») ou simplement des fichiers volumineux afin de saturer le stockage du serveur, ce qui entrave et nuit à la disponibilité du serveur
4. Remplacer un fichier existant sur le système
5. Contenu actif côté client (XSS, CSRF, etc.) pouvant mettre en danger d'autres utilisateurs si les fichiers sont accessibles au public.

Récupération publique de fichiers

Si le fichier téléchargé est accessible au public, d'autres menaces peuvent être prises en compte :

1. Divulgation publique d'autres fichiers
2. Lancer une attaque par déni de service en demandant un grand nombre de fichiers. Les demandes sont petites, mais les réponses sont beaucoup plus volumineuses.
3. Contenu de fichiers pouvant être considéré comme illégal, offensant ou dangereux (*par exemple,* données personnelles, données protégées par le droit d'auteur, etc.), ce qui fera de vous un hébergeur de fichiers malveillants.

Protection du téléchargement de fichiers

Il n'existe pas de solution miracle pour valider le contenu des utilisateurs. La mise en œuvre d'une approche de défense en profondeur est essentielle pour rendre le processus de téléchargement plus difficile et mieux adapté aux besoins et aux exigences du service. La mise en œuvre de plusieurs techniques est essentielle et recommandée, car aucune technique ne suffit à elle seule à sécuriser le service.

Validation des extensions

Assurez-vous que la validation a lieu après le décodage du nom de fichier et qu'un filtre approprié est mis en place afin d'éviter certains contournements connus, tels que les suivants :

- Les doubles extensions, *par exemple* `.jpg.php` , qui contournent facilement l'expre`\.jpg`égulière
- Octets nuls, *par exemple* `.php%00.jpg` , `.jpg` est tronquée et `.php` devient le nouveau où l'extension
- Expression régulière générique incorrecte qui n'a pas été correctement testée et vérifiée. Évitez de créer votre propre logique à moins d'avoir suffisamment de connaissances sur le sujet.

Reportez-vous à la validation des entrées CS pour analyser et traiter correctement l'extension.

Liste des extensions autorisées

Veillez à n'utiliser que les extensions *essentielles à l'activité*, sans autoriser aucun type d'extension *non requise*. Par exemple, si le système nécessite :

- le téléchargement d'images, autorisez un type convenu qui répond aux besoins de l'entreprise ;
- téléchargement de CV, autorisez `docx` et `pdf` extensions.

En fonction des besoins de l'application, veillez à utiliser les types de fichiers **les moins nuisibles** et présentant **le moins de risques**.

Bloquer les extensions

Identifiez les types de fichiers potentiellement dangereux et bloquez les extensions que vous considérez comme nuisibles pour votre service.

Veuillez noter que le blocage d'extensions spécifiques est une méthode de protection peu efficace en soi. L'article sur la vulnérabilité liée au téléchargement illimité de fichiers décrit comment les pirates peuvent tenter de contourner ce type de contrôle.

Validation du type de contenu

Le type de contenu des fichiers téléchargés est fourni par l'utilisateur et ne peut donc pas être considéré comme fiable, car il est facile à falsifier. Bien qu'il ne faille pas s'y fier pour la sécurité, il permet une vérification rapide afin d'empêcher les utilisateurs de télécharger involontairement des fichiers de type incorrect.

Outre la définition de l'extension du fichier téléchargé, son type MIME peut être vérifié pour une protection rapide contre les attaques simples par téléchargement de fichiers.

Cela peut être fait de préférence dans le cadre d'une approche par liste blanche ; sinon, cela peut être fait dans le cadre d'une approche par liste noire.

Validation de la signature du fichier

En conjonction avec la validation du type de contenu, la signature du fichier peut être vérifiée et comparée au fichier attendu qui devrait être reçu.

> Cette méthode ne doit pas être utilisée seule, car il est assez courant et facile de la contourner.

Sécurité des noms de fichiers

Les noms de fichiers peuvent mettre le système en danger de plusieurs façons, soit en utilisant des caractères non acceptables, soit en utilisant des noms de fichiers spéciaux et restreints. Pour Windows, consultez le guide MSDN suivant. Pour un aperçu plus complet des différents systèmes de fichiers et de la manière dont ils traitent les fichiers, consultez la page Wikipédia consacrée aux noms de fichiers.

Afin d'éviter la menace mentionnée ci-dessus, il est essentiel de créer une **chaîne aléatoire** comme nom de fichier, par exemple en générant un UUID/GUID. Si le nom de fichier est requis par les besoins de l'entreprise, une validation appropriée des entrées doit être effectuée pour les vecteurs d'attaque côté client (*par exemple,* contenu actif entraînant des attaques XSS et CSRF) et côté serveur (*par exemple,* écrasement ou création de fichiers spéciaux). Les limites de longueur des noms de fichiers doivent être prises en compte en fonction du système qui stocke les fichiers, car chaque système a sa propre limite de longueur de nom de fichier. Si les noms de fichiers des utilisateurs sont requis, envisagez de mettre en œuvre les mesures suivantes :

- Mettre en place une longueur maximale
- Limiter les caractères à un sous-ensemble autorisé, tel que les caractères alphanumériques, les tirets, les espaces et les points
  - Envisager d'indiquer à l'utilisateur ce qu'est un nom de fichier acceptable.
  - Limiter l'utilisation des points en début de nom (fichiers cachés) et des points consécutifs (traversée de répertoires).
  - Limitez l'utilisation d'un trait d'union ou d'espaces en début de ligne afin de sécuriser l'utilisation des scripts shell pour le traitement des fichiers.
  - Si cela n'est pas possible, bloquez les caractères dangereux qui pourraient compromettre le cadre et le système qui stockent et utilisent les fichiers.

Validation du contenu des fichiers

Comme mentionné dans la section Récupération de fichiers publics, le contenu des fichiers peut contenir des données malveillantes, inappropriées ou illégales.

En fonction du type attendu, une validation spéciale du contenu des fichiers peut être appliquée :

- Pour **les images**, l'application de techniques de réécriture d'images détruit tout type de contenu malveillant injecté dans une image ; cela peut être fait par randomisation.
- Pour **les documents Microsoft**, l'utilisation d'Apache POI permet de valider les documents téléchargés.
- **Les fichiers ZIP** ne sont pas recommandés, car ils peuvent contenir tous types de fichiers et les vecteurs d'attaque qui leur sont associés sont nombreux.

Le service de téléchargement de fichiers doit permettre aux utilisateurs de signaler les contenus illégaux et aux détenteurs de droits d'auteur de signaler les abus.

Si les ressources sont suffisantes, un examen manuel des fichiers devrait être effectué dans un environnement sandbox avant de les rendre publics.

Il peut être utile d'automatiser en partie la vérification, mais il s'agit d'un processus complexe qui doit être étudié de manière approfondie avant d'être utilisé. Certains services (*par exemple* Virus Total) fournissent des API permettant d'analyser les fichiers à la recherche de hachages de fichiers malveillants connus. Certains frameworks peuvent vérifier et valider le type de contenu brut

et le valider par rapport à des types de fichiers prédéfinis, comme dans la bibliothèque de dessins ASP.NET. Méfiez-vous des menaces de fuite de données et de la collecte d'informations par les services publics.

Emplacement de stockage des fichiers

L'emplacement où les fichiers doivent être stockés doit être choisi en fonction des exigences de sécurité et des besoins de l'entreprise. Les points suivants sont classés par ordre de priorité en matière de sécurité et sont exhaustifs :

1. Stockez les fichiers sur un **hôte différent**, ce qui permet une séparation complète des tâches entre l'application qui sert l'utilisateur et l'hôte qui gère les téléchargements et le stockage des fichiers.
2. Stockez les fichiers **en dehors du répertoire web**, où seul l'accès administratif est autorisé.
3. Stockez les fichiers **dans le répertoire racine du site Web** et attribuez-leur uniquement des autorisations d'écriture. Si un accès en lecture est nécessaire, il est indispensable de mettre en place des contrôles appropriés (*par exemple,* adresse IP interne, utilisateur autorisé, etc.).

Le stockage méthodique des fichiers dans des bases de données est une technique supplémentaire. Elle est parfois utilisée pour les processus de sauvegarde automatique, les attaques hors système de fichiers et les problèmes d'autorisations. En contrepartie, cela peut entraîner des problèmes de performances (dans certains cas), des considérations de stockage pour la base de données et ses sauvegardes, et cela ouvre la porte aux attaques SQLi. Cette technique n'est recommandée que lorsqu'un administrateur de base de données fait partie de l'équipe et que ce processus s'avère être une amélioration par rapport au stockage dans le système de fichiers.

> Certains fichiers sont envoyés par e-mail ou traités une fois téléchargés, et ne sont pas stockés sur le serveur. Il est essentiel de mettre en œuvre les mesures de sécurité décrites dans cette fiche avant d'effectuer toute action sur ces fichiers.

Autorisations utilisateur

Avant d'accéder à un service de téléchargement de fichiers, une validation appropriée doit être effectuée à deux niveaux pour l'utilisateur qui télécharge un fichier :

- Niveau d'authentification
  - L'utilisateur doit être un utilisateur enregistré ou identifiable afin de pouvoir définir des restrictions et des limitations pour ses capacités de téléchargement.
- Niveau d'autorisation
  - L'utilisateur doit disposer des autorisations appropriées pour accéder aux fichiers ou les modifier.

Autorisations du système de fichiers

> Définissez les autorisations des fichiers selon le principe du moindre privilège.

Les fichiers doivent être stockés de manière à garantir :

- Seuls les utilisateurs autorisés du système puissent lire les fichiers
- Seuls les modes requis soient définis pour le fichier
    - Si l'exécution est requise, il est recommandé, pour des raisons de sécurité, d'analyser le fichier avant de l'exécuter afin de s'assurer qu'il ne contient aucune macro ou script caché.

Limites de téléchargement et de chargement

L'application doit définir des limites de taille appropriées pour le service de téléchargement afin de protéger la capacité de stockage des fichiers. Si le système doit extraire les fichiers ou les traiter, la limite de taille des fichiers doit être prise en compte après la décompression des fichiers et en utilisant des méthodes sécurisées pour calculer la taille des fichiers zip. Pour plus d'informations à ce sujet, consultez la section Comment extraire en toute sécurité des fichiers à partir de ZipInputStream, le flux d'entrée Java permettant de gérer les fichiers ZIP.

L'application doit également définir des limites de requêtes appropriées pour le service de téléchargement, si disponible, afin de protéger le serveur contre les attaques par déni de service.

Extraits de code Java

Référentiel de protection du téléchargement de documents écrit par Dominique pour certains types de documents en Java.