

Préparer le nid

Avant que le CasaOS ne puisse rapporter ses trouvailles, il lui faut un endroit où les déposer. Dans **Paperless-ngx**, cet endroit s'appelle le dossier "**Consume**". C'est un dossier que Paperless surveille 24h/24 : dès qu'un fichier tombe dedans, il l'attrape, l'analyse et le range.

Créer un point de chute

Dans mon installation, le dossier "**Consume**" n'est pas sur le Mac Mini, ni sur le serveur de scan. Il se situe dans un NAS.

J'ai fait ce choix pour que le dossier consume soit accessible facilement dans un point de stockage déjà connu de mes machines.

Sur mon NAS, j'ai créé un dossier "Aspirateur_de_documents" qui sera mon dossier consume.



J'ai ensuite monté ce dossier du NAS dans un répertoire local (`/mnt/aspirateur_de_documents`) sur mon **CasaOS** grâce au fichier `/etc/fstab`, comme ça en plus, il se monte automatiquement au démarrage.

J'ai fait également la même chose sur le **Mac Mini** qui héberge **Paperless-ngx** pour qu'il surveille ce dossier.

Le réglage crucial ici, c'est le protocole de partage, le dossier est partagé en SMB, le serveur **CasaOS** a les droits en écriture et le **Mac Mini** lui a les droits en écriture, mais aussi en suppression, sinon il ne pourrait pas supprimer le fichier qu'il a importé et essaierait de l'importer en boucle (le pauvre).

Configuration de Paperless-ngx

C'est ici, sur mon **Mac Mini**, que réside l'intelligence du système, Paperless-ngx est prêt, mais il faut lui expliquer où se trouve sa "gamelle" (le fameux dossier **Aspirateur_de_documents** sur le NAS) et comment il doit se comporter.

Le branchement vers Docker

Comme la plupart des installations autohébergées, mon Paperless tourne dans un container Docker

Pour qu'il puisse "voir" le dossier du NAS, j'ai dû créer un pont dans mon fichier de configuration (`docker-compose.yml`).

Dans la section `volumes`, j'ai ajouté cette ligne :

- `/Volumes/aspirateur_de_documents:/usr/src/paperless/consume`

C'est comme si je disais à Paperless : "Ce dossier que tu vois sur le réseau, considère-le comme ton dossier d'entrée".

Les ordres de mission (Variables d'environnement)

Pour que le traitement soit aux petits oignons, j'ai ajusté quelques réglages dans le fichier de configuration de Paperless:

`PAPERLESS_CONSUMER_POLLING=60` : C'est mon réglage préféré. Comme le dossier est sur un NAS (en réseau), les notifications automatiques de fichiers ne marchent pas toujours bien. En mettant "60", Paperless va toquer à la porte du NAS toutes les minutes pour voir s'il y a du nouveau. C'est le secret de la réactivité !

`PAPERLESS_OCR_LANGUAGE=fra` : On n'oublie pas de lui dire de lire le français, sinon il va bégayer sur nos accents.

La magie du tri automatique

C'est là que Paperless devient incroyable. Une fois qu'il a "aspiré" le PDF envoyé par le serveur CasaOS (ou mis dans le NAS manuellement), il ne se contente pas de le stocker.



1. **Il va lancer un OCR (Reconnaissance Optique de Caractères)** : Il transforme l'image du scan en texte recherchable.
2. **Il analyse** : S'il voit des mots comme "Facture" et "EDF", il va appliquer tout seul les tags que j'ai pré-configuré (Par exemple le tag "Facture", "Appartement", "Électricité"...)
3. **Il archive** : Il va renommer mon fichier proprement et le ranger dans sa base de données

À ce stade, toute la logistique de réception est en place. Le **NAS** sert de boîte aux lettres intelligente et **Paperless-ngx** guette patiemment, prêt à bondir sur le moindre PDF qui se trouvera dans le dossier.

Mais pour l'instant, notre "Aspirateur de documents" est bien vide...

Il est temps de passer à l'action et de voir comment on va faire pour envoyer des documents dans ce dossier.

Révision #2

Créé 2026-04-05 19:47:27 UTC par Renard

Mis à jour 2026-04-05 20:58:47 UTC par Renard