

1. RÉVISIONS SUR LES REGEX

Pour les révisions, nous allons utiliser des outils en ligne pour apprendre à manipuler les regex avec des exemples concrets et sophistication progressive des regex que nous allons construire.

1. Rendez-vous sur ce site : <https://regexone.com/> et effectuez les exercices
2. Une fois terminé, faites les challenges de ce site en allant sur « Interactive Tutorial » en haut à droite puis dans la section « Practice problems » faites les problèmes 1 à 3

2. CAS PRATIQUE SUR LE FICHIER ALLOCLIST.TXT

Téléchargez le fichier texte contenant toutes les allocations du RIPE : <https://ftp.ripe.net/ripe/stats/membership/alloclist.txt>

Pour cette partie, nous allons utiliser l'outil d'expression régulière disponible dans VSCode. Ouvrez le fichier texte avec VSCode (ou VSCodium), lancez la commande CTRL+F et cochez la case « Use Regular Expression » en forme de carré et d'étoile (ou ALT+R)

2.1. RECHERCHE DANS LE FICHIER

Le but ici sera de rechercher des informations précises dans ce fichier. A chaque fois, on pourra utiliser le bouton « suivant » pour lancer une recherche dans le fichier en grâce aux regex demandées. Proposez des regex pour rechercher :

1. Une IPv4 (sans faire de vérification sur la validité de chaque octet)
2. Une IPv4 dont le premier octet est 212
3. Un bloc d'IPv4 attribué en 2019

2.2. UTILISER LES REGEX POUR GÉNÉRER UNE CONFIGURATION

Toujours dans le fichier des allocations du RIPE, on souhaite conserver uniquement les IPv4. Pour cela :

1. faire un CTRL+F et renseigner une regex qui sélectionne les **lignes complètes** qui contiennent des IPv4 (voir les caractères spéciaux pour début et fin de ligne en regex)
2. si la regex est bonne, utiliser le raccourci CTRL+SHIFT+L pour sélectionner toutes les listes dans l'éditeur
3. faire un CTRL+C, ouvrir un deuxième fichier et faire CTRL+V

A partir de là, on n'a que les lignes qui contiennent des IPv4. On souhaite par la suite à retravailler ce fichier pour générer la configuration à rentrer dans un filtre BGP (ip prefix-list) pour n'autoriser que les IP affectées par le RIPE. Pour chaque ligne du fichier texte contenant un sous-réseau IPv4, nous devons générer une ligne de configuration du type :

```
ip prefix-list PL-RIPE permit X.X.X.X/NN le 24
```

Écrire une regex permettant de rechercher un sous-réseau IPv4 **et** son masque :

1. Reprendre ou écrire la regex qui permet de sélectionner les IPv4 et les masques sur chaque ligne
2. Placer des parenthèses afin de créer un atome pour le numéro de réseau, et un atome pour le masque

4. Dans le champ « Replace », proposer un texte de remplacement permettant de remplacer chaque ligne par une ligne ip prefix-list correspondant au sous-réseau de la ligne. (Utilisez les variable 1, 2, etc.)

3. UTILISATION DES REGEX DANS UN SCRIPT PYTHON

Maintenant, nous allons utiliser les expressions régulières dans un script Python. Cela va nous permettre d'enrichir ce que nous avons fait précédemment.

En effet, juste avant, nous avons créé une prefix-list à partir du fichier d'allocation des IP du RIPE mais nous n'avons pas spécifié le numéro d'ordre. Qui plus est, il reste encore plein de lignes parasites dans notre fichier retravaillé. Nous allons donc passer par un script Python.

Téléchargez le script td2.py dans le même dossier que le fichier alloclist.

Ouvrez le script et mettez votre regex trouvée à la question 6 de l'exercice précédent dans la variable REGEX_IPV4. N'enlevez pas le r devant les guillemets.

Lancez le script avec l'option -f pour spécifier le chemin du fichier alloclist.txt et -4. Vous devriez voir apparaître la prefix-list sans ligne parasite et avec le numéro d'ordre. Avec l'option -6, vous verrez la même chose mais en IPv6. Maintenant, regardez plus en détail le code.

- Quel module Python a été utilisé pour les regex ?
- D'après la documentation de ce module sur Internet, pourquoi avons-nous utilisé des raw string (chaînes de caractères commençant par r) ?
- Décrivez le fonctionnement de la fonction prefix_list4 : qu'est-ce que la fonction prend comme paramètre d'entrée, ce qu'elle effectue, en particulier la méthode utilisée pour rechercher une expression régulière dans une chaîne de caractères avec et ce qu'elle renvoie