Gestion des index

NOTES pour ce chapitre

Pour ce chapitre, il faut avoir démarré une instance de Kibana connectée Elasticsearch.

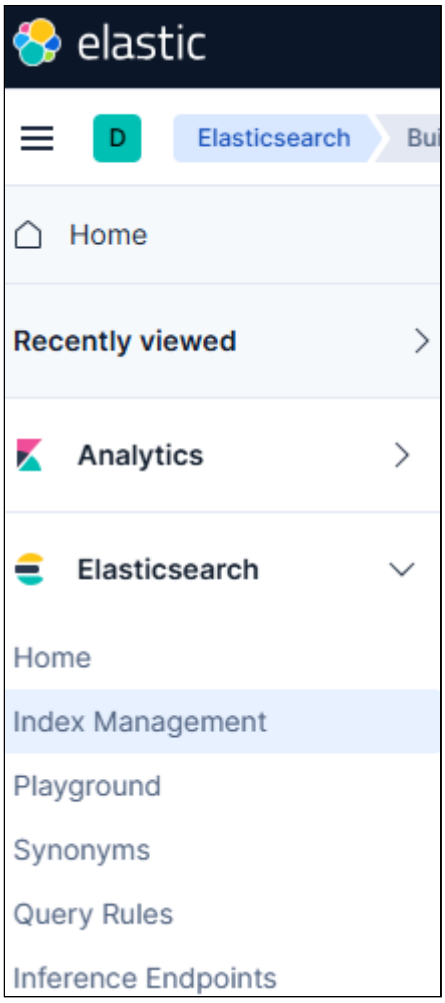
Permissions requises

Si vous utilisez les fonctionnalités de sécurité Elasticsearch, les privilèges de sécurité suivants sont requis :

- Le privilège cluster `monitor` pour accéder aux fonctionnalités de gestion des index de Kibana
- Les privilèges d'index `view_index_metadata` et `manage` pour visualiser les données d'un flux de données ou d'un index
- Le privilège cluster `manage_index_templates` pour gérer les modèles d'index

Pour ajouter ces privilèges, accédez à **Stack Management > Security > Roles** ou utilisez l'API **Create or update roles**.

Examinez vos index et effectuez des opérations depuis la vue **Indices** en allant sur la page : Index Management - Elastic



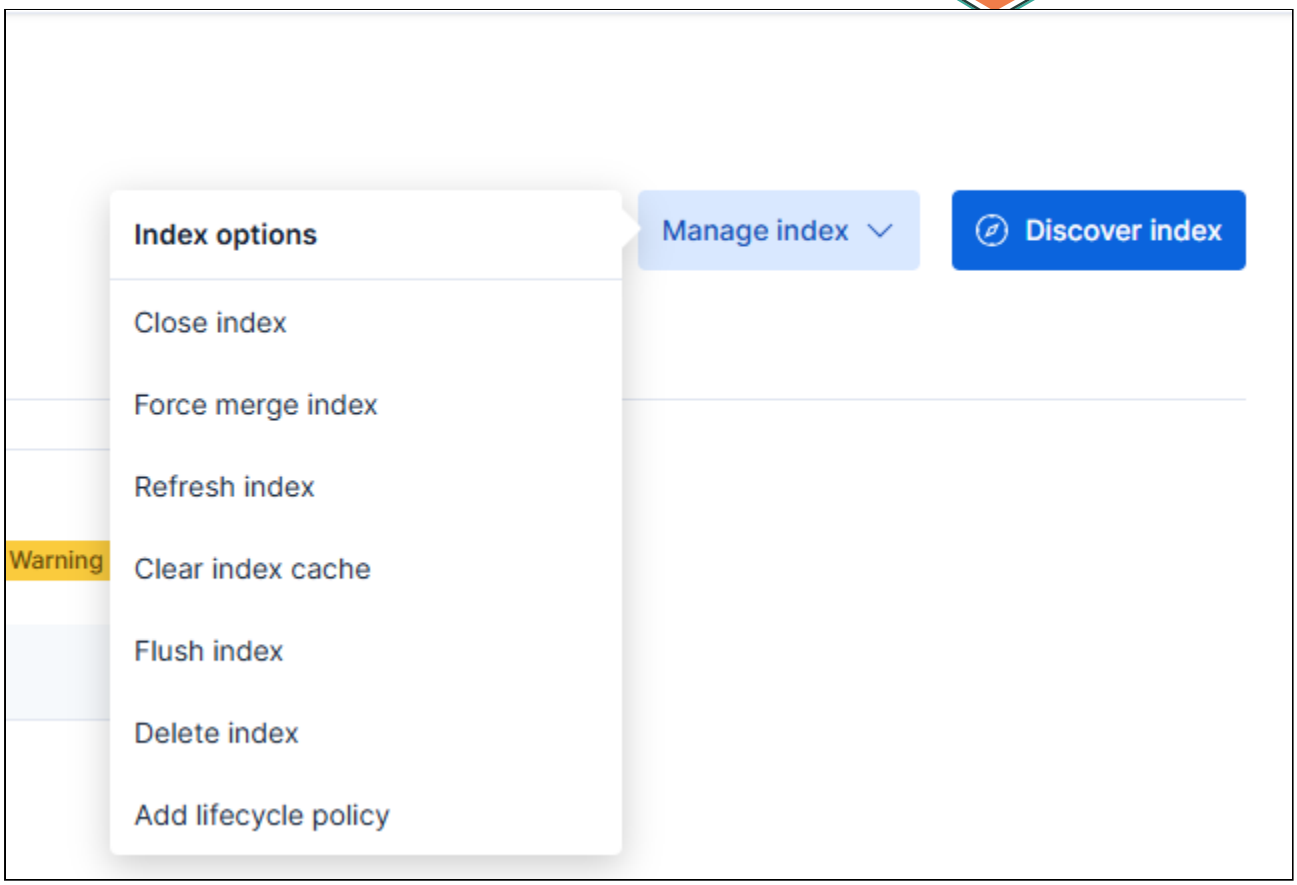
Actions disponibles

Pour un index unique :

- Cliquez sur le nom de l'index pour accéder aux détails et effectuer des opérations
- Accédez à la vue d'ensemble de l'index, aux mappages et aux paramètres
- Depuis cette vue, vous pouvez naviguer vers **Discover** pour explorer davantage les documents de l'index

Pour plusieurs index :

- Sélectionnez leurs cases à cocher
- Ouvrez le menu **Manage** (Gérer)
- Effectuez des opérations telles que **close**, **forcemerge** et **flush** (sur Elastic Stack)



---

1. Close — fermer un index

Désactive un index pour libérer des ressources.
Les données restent sur disque, mais l'index n'est plus accessible tant qu'il n'est pas rouvert.

```
POST /mon_index/_close
```

Pour le rouvrir :

```
POST /mon_index/_open
```

Usage : utile pour archiver temporairement un index non utilisé.
Effet : l'index ne consomme plus de mémoire (heap).

---

2. Forcemerge — fusionner les segments

Elasticsearch stocke les données d'un index dans plusieurs segments (petits fichiers sur le disque).
Chaque fois que vous ajoutez, modifiez ou supprimez un document, un nouveau segment est créé.

Avec le temps, ces segments s'accumulent et ralentissent les recherches.
La commande **forcemerge** regroupe plusieurs segments en un seul plus gros.

```
POST /mon_index/_forcemerge?max_num_segments=1
```

forcemerge permet :

- d'accélérer les recherches,
- de réduire l'espace disque utilisé,
- mais au prix d'une forte charge temporaire sur le serveur pendant l'opération.

3. Flush — forcer l'écriture des données sur disque

Le flush force Elasticsearch à enregistrer sur le disque toutes les écritures encore en mémoire.

Quand un document est ajouté ou modifié, il est d'abord gardé dans un tampon temporaire appelé translog.
Le flush vide ce tampon et enregistre ces données de façon durable sur le disque.

```
POST /mon_index/_flush
```

Flush permet de :

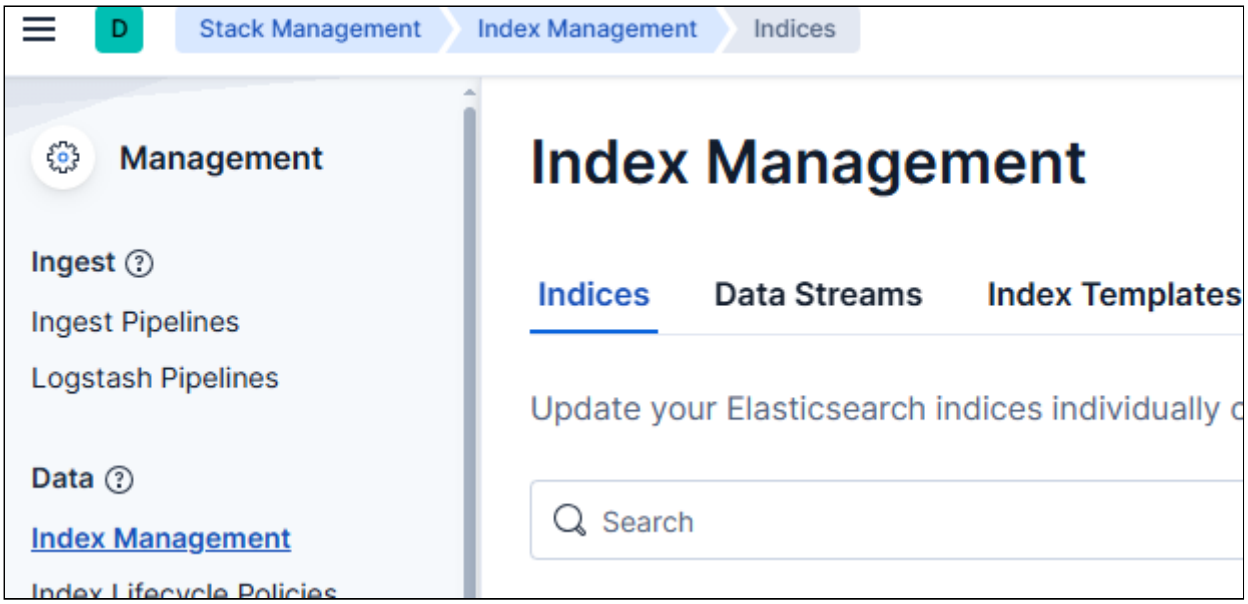
- sécuriser les données (elles ne sont plus perdues en cas d'arrêt brutal),
- réduire la taille du translog,
- et stabiliser l'index avant une sauvegarde ou un redémarrage.

Résumé :

Opération	Fonction principale	Risque/Coût	Utilisation typique
`close`	Désactiver un index	Faible	Archiver temporairement
`forcemerge`	Compacter les segments	Élevé (CPU, I/O)	Après nettoyage ou migration
`flush`	Écrire en dur les données	Faible	Avant sauvegarde ou arrêt du cluster

Options d'affichage et filtrage

Dans le menu management → Index Management



- Activez **Include hidden indices** pour afficher l'ensemble complet des index, y compris les index de support pour les flux de données

- Utilisez la barre de recherche ou cliquez sur un badge pour filtrer la liste des index
- Les badges indiquent si un index est un **index suiveur (follower index)**, un **index de cumul (rollup index)** ou **gelé (frozen)**

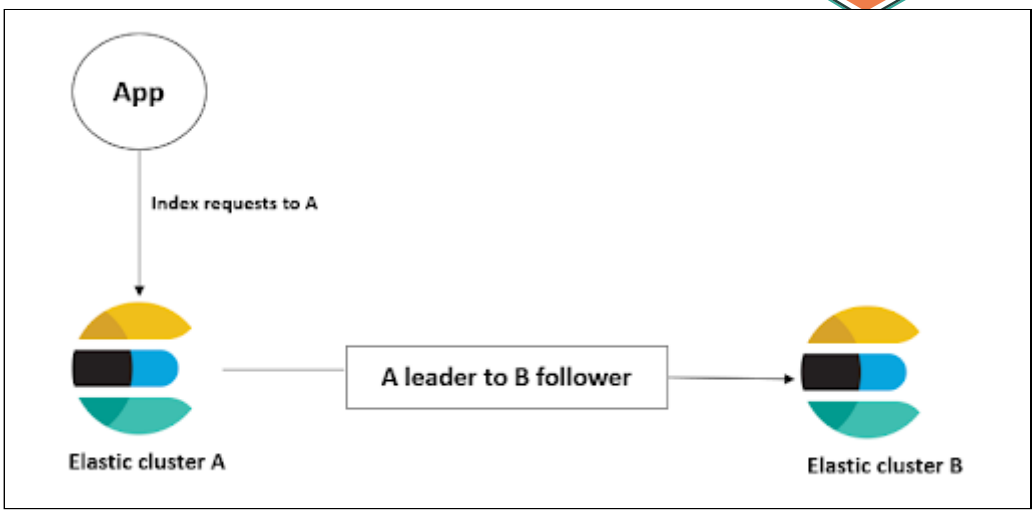
Types d'index

Index suiveur (follower index)

Copie en temps réel un **index principal (leader)** d'un autre cluster via la **Cross-Cluster Replication (CCR)**.

- Sert à **répliquer les données** pour la haute disponibilité ou la lecture distante.
- Se met automatiquement à jour quand le leader change.

Exemple :
Un cluster de production alimente un cluster de secours en lecture seule.

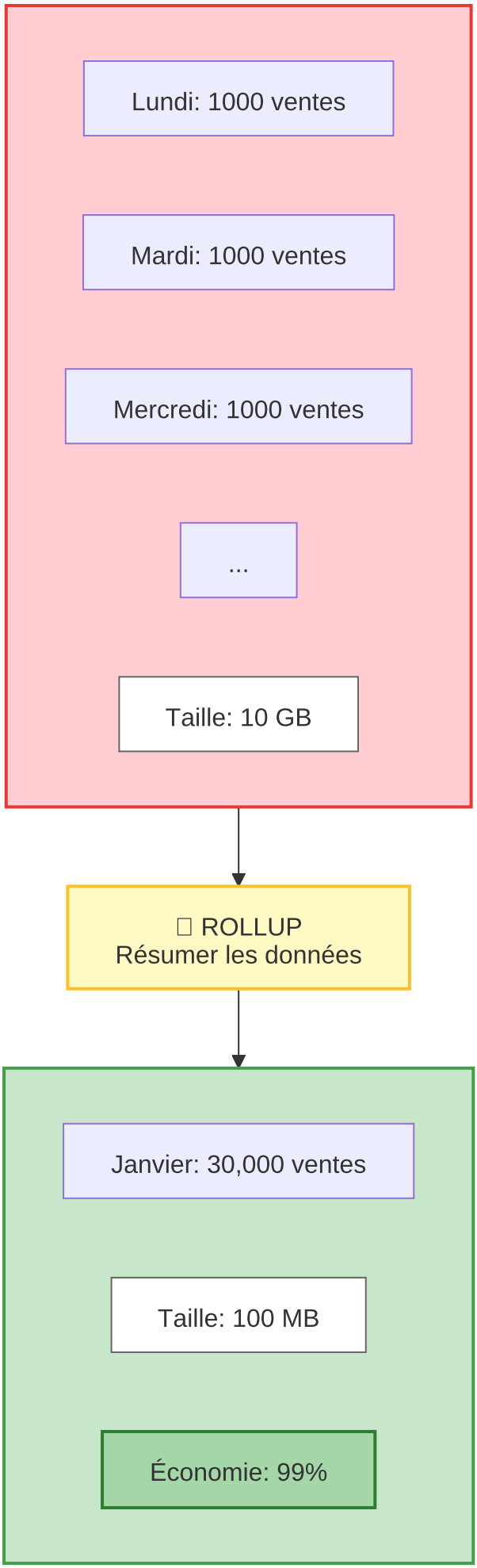


Index de cumul (rollup index)

Contient des **données agrégées** à partir d'un index source (ex. moyennes, totaux, statistiques).

- Réduit la taille des données historiques.
- Optimise les requêtes analytiques sur de longues périodes.

Exemple :
Remplacer des millions de logs journaliers par des moyennes horaires.


Lundi: 1000 ventes
Mardi: 1000 ventes
Mercredi: 1000 ventes
...
Taille: 10 GB
 ROLLUP
Résumer les données
Janvier: 30,000 ventes
Taille: 100 MB
Économie: 99%

Index gelé (frozen index)

Index archivé en **lecture seule** et **chargé à la demande** depuis le disque.

- Réduit fortement la mémoire utilisée.
- Plus lent à interroger (car doit être rechargé en mémoire).

Exemple :
Garder plusieurs années de logs accessibles sans consommer de ressources.

Résumé :

Type d'index	Fonction principale	Lecture	Écriture	Objectif
Follower	Réplication d'un index distant	Oui	Non	Haute disponibilité
Rollup	Données agrégées et compactées	Oui	Non	Analyse historique
Frozen	Données archivées, chargées à la demande	Oui (lent)	Non	Archivage économique

- **Sur Elastic Stack** : Vous pouvez également utiliser les menus déroulants pour filtrer la liste par statut ou phase du **cycle de vie de l'index (Index Lifecycle Management)**