

Architectures de recherche des sites Internet pour SharePoint Server 2016

Vue d'ensemble

Ce modèle illustre une architecture de batterie de serveurs pour sites Internet de taille moyenne. Les conditions d'architecture requises peuvent varier en fonction de la composition des données analysées (taille des éléments et formats). Cet exemple illustre le type des composants de recherche nécessaires et leur nombre respectif. Utilisez cet exemple comme point de départ pour préparer vos propres environnements de recherche. Pour plus d'informations sur les processus de recherche et les interactions entre les composants de recherche, voir Architectures de recherche pour SharePoint® Server 2016 (<https://go.microsoft.com/fwlink/p/?linkid=832554>).

Composants de recherche

Index

Composant d'index

Le composant d'index est la représentation logique d'une copie d'index.

Partitions d'index

- Vous pouvez diviser l'index en partitions discrètes, chacune contenant une partie distincte de l'index.
- Une partition d'index est stockée dans un ensemble de fichiers sur un disque.
- L'index de recherche est le regroupement de toutes les partitions d'index.

Copies d'index

- Chaque partition d'index comporte une ou plusieurs copies d'index contenant les mêmes informations.
- Vous devez fournir un composant d'index pour chaque copie d'index.
- Pour obtenir une tolérance de panne et une redondance, créez des copies d'index supplémentaires pour chaque partition d'index et distribuez-les sur plusieurs serveurs d'applications.

Traitement des requêtes

Composant de traitement des requêtes

Permet d'analyser et de traiter les requêtes et les résultats de la recherche.

Administrateur

Composant d'administration de la recherche

Permet d'exécuter des processus système obligatoires pour la recherche. Il peut y avoir plusieurs composants d'administration de la recherche par application Service de recherche, mais un seul à la fois peut être actif.

Analyse

Composant d'analyse

Permet d'analyser le contenu en fonction de ce qui est indiqué dans la base de données d'analyse.

Traitement de contenu

Composant de traitement de contenu

Permet d'exécuter plusieurs processus sur les éléments analysés, tels que l'analyse syntaxique et le mappage de propriété.

Analyse

Composant de traitement de l'analyse

Permet d'exécuter l'analyse de la recherche et l'analyse de l'utilisation.

Bases de données de recherche

Base de données des paramètres de recherche

Base de données des paramètres de recherche

Permet de stocker les données de configuration de recherche. Il n'existe qu'une seule base de données des paramètres de recherche par application Service de recherche.

Base de données de recherche (robot)

Base de données d'analyse

Stocke l'historique des analyses et gère les opérations d'analyse. Chaque base de données d'analyse peut être associée à un ou plusieurs composants d'analyse.

Base de données de liens

Base de données de liens

Stocke les informations extraites par le composant de traitement de contenu et également les informations de consultation.

Base de données d'analyse

Base de données de création de rapports d'analyse

Permet de stocker les résultats de l'analyse de l'utilisation.

Considérations liées aux performances

Serveurs d'applications hébergeant des composants de recherche

INFORMATIONS	RAISONNEMENT
Cache	La requête et ses résultats sont mis en cache avec Windows Server AppFabric dans des paires clé/valeur : la requête correspond à la clé et les résultats correspondent à la valeur. Pour chaque requête, il existe un rapport de cache approximatif de 50 %. Cela signifie que, si vous avez un modèle d'utilisation de 200 requêtes par seconde, environ 100 requêtes seront envoyées à l'index de recherche et les 100 autres requêtes seront mises en cache. Les résultats du cache ont une latence de requête inférieure à ceux de l'index de recherche. Par exemple, les résultats des requêtes de première page, dont l'exécution est répétée, sont susceptibles d'être mis en cache.
Analyse continue	Nous vous recommandons d'activer l'analyse continue avec un intervalle d'une minute au lieu de l'intervalle par défaut de 15 minutes. Vous ne pouvez activer l'analyse continue que sur les sources de contenu SharePoint.
Accès anonyme	Avec un accès anonyme, les utilisateurs n'ont pas à utiliser des informations d'identification pour se connecter à un site Internet SharePoint. En outre, les requêtes anonymes sont mises en cache. Elles s'avèrent donc moins chères en raison d'une latence de requête plus faible. Vous devez activer l'accès anonyme à deux emplacements : sur le site web frontal et sur le site lui-même.
Latence de requête	La latence des requêtes est influencée par la mise en cache, l'accès anonyme et d'autres facteurs, tels que le nombre et la complexité des règles de requête appliquées et déclenchées. Prenez également en compte les disques sur lesquels l'index de recherche est stocké. Un disque qui comporte plusieurs broches peut améliorer la vitesse d'accès du disque et réduire la latence des requêtes.

Topologies de recherche pour des sites Internet

Les topologies pour les sites Internet présentent les différences suivantes avec les autres topologies de recherche :

- Les serveurs web sont incorporés dans la topologie pour sites Internet afin de traiter les demandes des utilisateurs. Il n'en est pas de même dans les topologies de recherche de contenu d'entreprise, où les serveurs web ne font pas partie de la batterie de serveurs de recherche dédiée.
- Le composant de traitement des requêtes et les composants d'index résident sur des serveurs web pour une utilisation maximale des ressources matérielles disponibles et une simplification de la mise à l'échelle de la topologie de recherche.

Configuration matérielle requise

Cette configuration requise s'applique à la topologie de recherche pour sites Internet de taille moyenne. La quantité minimale de mémoire vive (RAM) requise indiquée pour un serveur qui héberge un composant de recherche correspond à la quantité totale de mémoire vive (RAM) requise pour ce serveur. Par exemple, si vous hébergez un composant de traitement de contenu, un composant d'administration de recherche et un composant d'analyse sur un serveur, la quantité totale de mémoire vive (RAM) minimale requise pour ce serveur est de 24 Go.

Configuration matérielle minimale requise pour les serveurs d'applications

COMPOSANT DE RECHERCHE SUR LE SERVEUR PHYSIQUE	Mémoire RAM	DISQUE DUR	PROCESSEUR
Composant d'index	48 Go pour chaque serveur de la batterie de serveurs qui héberge un composant d'index, un composant de traitement des requêtes et le serveur web frontal.	500 Go d'espace disque supplémentaire, de préférence un volume/une partition de disque distinct.	TOUS LES COMPOSANTS : 64 bits, 4 cœurs minimum, 8 cœurs recommandés.
Composant de traitement de l'analyse	24 Go pour chaque serveur de la batterie qui héberge un composant de traitement de l'analyse, un composant d'analyse et/ou un composant d'administration de recherche.	300 Go d'espace disque supplémentaire, de préférence un volume/une partition de disque distinct. Cet espace disque est nécessaire pour le traitement local des données d'analyse avant que celles-ci soient écrites dans la base de données de création de rapports d'analyse.	
Composant d'analyse/Composant de traitement de contenu	Voir la configuration requise pour le composant de gestion de l'analyse.	80 Go pour le lecteur système. Vous devez disposer d'une quantité d'espace suffisante pour l'installation de base et pour les diagnostics (tels que la journalisation et le débogage), pour la création d'images mémoire et pour d'autres opérations. Pour une utilisation en mode production, vous devez également disposer d'un espace disque supplémentaire pour les opérations habituelles. La quantité d'espace libre nécessaire doit représenter le double de la RAM disponible pour les environnements de production.	
Composant de traitement des requêtes	Voir la configuration requise pour le composant d'index.		
Composant d'administration de la recherche	Voir la configuration requise pour le composant de gestion de l'analyse.		

Configuration matérielle minimale requise pour les serveurs de base de données

COMPOSANT MATÉRIEL	CONFIGURATION REQUISE
Processeur	64 bits, 4 cœurs pour les déploiements
Mémoire RAM	8 Go pour les petits déploiements
Disque dur	16 Go pour les déploiements moyens
	80Go pour le lecteur système

Batterie de serveurs de sites Internet de taille moyenne (environ 85 pages consultées par seconde)

Cette batterie de serveurs est conçue pour fournir une topologie de recherche SharePoint Server 2016 à tolérance de panne qui est optimisée pour un corpus de 3 400 000 éléments. La batterie de serveurs fournie en exemple traite entre 100 et 200 documents par seconde, en fonction de la langue, et permet 85 consultations de page par seconde et 100 requêtes par seconde.

Cette topologie est optimisée pour le matériel physique, mais vous pouvez également la déployer sur des machines virtuelles.

