Architetture di ricerca
per SharePoint 2016

Panoramica

L'architettura della ricerca in Microsoft® SharePoint® Server 2016 è costituita da componenti e database che operano in modo coesivo per eseguire l'operazione di ricerca. Tutti i componenti risiedono nei server dell'applicazione e tutti i database risiedono sui server del database.

Interazione dei componenti di ricerca

Processi di ricerca per indicizzazione e dei componenti

La ricerca per indicizzazione e l'architettura di elaborazione del contenuto include il **componente di ricerca per indicizzazione**, **il database di ricerca per indicizzazione** e **il componente di elaborazione del contenuto**. Per entrambi i componenti di ricerca è possibile eseguire la scalabilità orizzontale in base alle esigenze di volume della ricerca per indicizzazione e ai requisiti delle prestazioni.

1 Informazioni sul **componente di ricerca per indicizzazione**

- Il componente di ricerca per indicizzazione è responsabile per le origini dei contenuti di ricerca per indicizzazione. Offre elementi sottoposti a ricerca per indicizzazione, sia sul contenuto effettivo sia sui metadati associati, per il componente di elaborazione del contenuto.
- Il componente di ricerca per indicizzazione richiama i connettori o gestori di protocolli che interagiscono con le origini di contenuto per recuperare i dati. È possibile distribuire contemporaneamente più componenti di ricerca per indicizzazione per la ricerca per indicizzazione.
- Il componente di ricerca per indicizzazione utilizza uno o più database di ricerca per indicizzazione per memorizzare temporaneamente informazioni sugli elementi di ricerca per indicizzazione e per tenere traccia della cronologia di ricerca per indicizzazione.

A Informazioni sul **database di ricerca per indicizzazione**

- Il database di ricerca per indicizzazione contiene l'analisi dettagliata e le informazioni cronologiche sugli elementi sottoposti a ricerca per indicizzazione.
- Il database mantiene informazioni quali l'ora dell'ultima ricerca per indicizzazione, l'ultimo ID di ricerca per indicizzazione e il tipo di aggiornamento effettuato durante l'ultima ricerca per indicizzazione.

2 Informazioni sul **componente di elaborazione del contenuto**

- Il componente di elaborazione del contenuto è compreso tra il componente di ricerca per indicizzazione e il componente di indicizzazione. Elabora gli elementi sottoposti a ricerca e li inserisce nel componente di indicizzazione.
- Il componente di elaborazione del contenuto trasforma gli elementi sottoposti a ricerca per indicizzazione in elementi che possono essere inclusi nell'indice di ricerca eseguendo operazioni quali l'analisi del documento e il mapping delle proprietà.
- Il componente di elaborazione del contenuto e il componente di elaborazione di query eseguono elaborazioni linguistiche. Esempi di elaborazione linguistica durante l'elaborazione del contenuto includono il rilevamento della lingua e l'estrazione di entità.
- Il componente di elaborazione del contenuto scrive le informazioni sui collegamenti e URL nel database di collegamento.

Processi di indicizzazione e query

L'architettura di indicizzazione e query include il **componente di indicizzazione**, **la partizione di indice** e **il componente di elaborazione di query**, tutti elementi scalabili in base al volume del contenuto, volume della query e requisiti delle prestazioni.

4 Informazioni sul **componente di indicizzazione**

- Un componente di indicizzazione è la rappresentazione logica di una replica dell'indice. Nell'architettura di ricerca è necessario eseguire il provisioning di un componente di indicizzazione per ogni replica dell'indice.
- Il componente di indicizzazione riceve gli elementi elaborati dal componente di elaborazione del contenuto e li scrive nel file di indice.
- Il componente di indicizzazione riceve le query dal componente di elaborazione di query e restituisce i set di risultati.
- Le query vengono inviate alle repliche di indice tramite il componente di elaborazione di query. Il sistema indirizza e il carico bilancia le query in arrivo con le repliche di indice.

Informazioni **sulla partizione di indice**

- Una partizione di indice è una parte logica dell'indice di ricerca intero. L'indice di ricerca è l'aggregazione di tutte le partizioni di indice.

5 Informazioni sul **componente di elaborazione di query**

- Il componente di elaborazione di query è compreso tra la ricerca front-end e il componente di indicizzazione.
- Il componente di elaborazione di query analizza ed elabora le query di ricerca e i risultati.
- Il componente di elaborazione di query e il componente di elaborazione del contenuto eseguono elaborazioni linguistiche. Esempi di elaborazione linguistica durante l'elaborazione di query includono interruzioni di parole-e stemming.
- Quando il componente di elaborazione di query riceve una query dalla ricerca front-end, analizza ed elabora le query per tentare di ottimizzare precisione, richiamo e pertinenza. La query elaborata viene quindi inviata al componente di indicizzazione.
- Il componente di indicizzazione restituisce un set di risultati in base alla query elaborata al componente di elaborazione di query, che in risposta elabora il set di risultati prima di inviarlo nuovamente al front-end di ricerca.

Amministrazione della ricerca

L'amministrazione della ricerca è costituita dal **componente di amministrazione della ricerca** e il database corrispondente.

6 Informazioni sul **componente di amministrazione della ricerca**

- Il componente di amministrazione della ricerca è responsabile dell'esecuzione di un numero di processi di sistema essenziali per la ricerca.
- Questo componente esegue il provisioning, che consiste nell'aggiunta e nell'inizializzazione di ulteriori istanze dei componenti di ricerca.

D Informazioni sul **database di amministrazione della ricerca**

- Il database di amministrazione della ricerca memorizza i dati di configurazione della ricerca, ad esempio la topologia, le regole di ricerca per indicizzazione, le regole di query e i mapping tra proprietà sottoposte a ricerca per indicizzazione e proprietà gestite.

Ruoli di server

Server Web

- Web part di ricerca host e pagine Web part per rispondere alle query di ricerca.
- Nelle farm di ricerca dedicate per ricerca contenuti organizzazione, questo ruolo non è necessario poiché i server Web presso le farm remote contattano i server che servono direttamente le query.
- Questo ruolo è necessario per le farm che includono altre funzionalità di SharePoint Server 2016.
- Nelle piccole farm, questo ruolo può essere condiviso con il ruolo del server applicazioni.

Server applicazioni con componenti di ricerca

- Ospita tutti i componenti di ricerca se solo un server è configurato. In caso contrario, contiene componenti associati al server, come configurato dall'amministratore.
- Contiene l'intero indice di ricerca se solo una partizione di indice è configurata. In caso contrario, contiene parti dell'indice associate a partizioni dell'indice come configurato dall'amministratore.
  - ❑ Il componente di elaborazione di query indirizza le query in arrivo alle repliche di indice.
  - ❑ Ogni replica di indice è un componente di indicizzazione.
  - ❑ Almeno una partizione di indice deve essere configurata per ogni farm.
- Almeno un componente della ricerca deve essere configurato per ogni farm.
- In un server applicazioni non è possibile avere più componenti di ricerca dello stesso tipo.
- Aggiungere i componenti di ricerca in altri server per garantire la ridondanza.

Server di database

- Ospita database di ricerca.
- Può ospitare altri database di SharePoint Server 2016.
- Può essere speculare o in cluster.
- Per migliorare prestazioni e capacità, prendere in considerazione l'aggiunta di dischi al server di database o server di database (a seconda dei colli di bottiglia).

Componenti di ricerca

Nessuno

In SharePoint Server 2016 per architetture di ricerca contenuti organizzazione, i componenti di ricerca non sono contenuti nei server Web.

Non è il caso delle architetture di ricerca di siti Internet. Qui, il componente di elaborazione query e i componenti dell'indice risiedono sui server Web per utilizzare al massimo le risorse hardware disponibili e semplificare la scalabilità orizzontale della topologia di ricerca.

Indice

Componente di indicizzazione
Il componente di indicizzazione è la rappresentazione logica di una replica dell'indice.

Partizioni di indice
- È possibile dividere l'indice in porzioni discrete, ciascuna contenente una parte distinta dell'indice.
- Una partizione dell'indice viene archiviata in una serie di file su un disco.
- L'indice di ricerca è l'aggregazione di tutte le partizioni di indice.

Repliche di indice
- Ogni partizione dell'indice contiene una o più repliche di indice con le stesse informazioni.
- È necessario eseguire il provisioning di un componente di indicizzazione per ogni replica di indice.
- Per ottenere la tolleranza di errore e la ridondanza, creare ulteriori repliche di indice per ogni partizione dell'indice e distribuirle su più server applicazioni.

Elaborazione query

Componente di elaborazione di query
Analizza ed elabora le query di ricerca e i risultati.

Admin

Componente di amministrazione della ricerca
- Esegue processi di sistema essenziali per la ricerca.
- Possono essere presenti più componenti di amministrazione della ricerca per Applicazione servizio di ricerca, ma può essere attivo un solo componente alla volta.

Ricerca per indicizzazione

Componente di ricerca per indicizzazione
Esegue la ricerca per indicizzazione del contenuto in base a quanto specificato nei database di ricerca per indicizzazione.

Elaborazione contenuto

Componente di elaborazione del contenuto
Esegue diversi processi sulle voci sottoposte a ricerca per indicizzazione, ad esempio l'analisi dei documenti e il mapping delle proprietà.

Analisi

Componente di elaborazione dei dati analitici
Esegue l'analisi della ricerca e dell'utilizzo.

DB di amministrazione ricerca

Database di amministrazione della ricerca
Archivia i dati di configurazione della ricerca. Per ogni applicazione di servizio di ricerca è consentito un solo database di amministrazione delle ricerche.

DB di ricerca per indicizzazione

Database di ricerca per indicizzazione
- Memorizza la cronologia della ricerca per indicizzazione.
- Gestisce le operazioni di ricerca per indicizzazione.
- A ogni database di ricerca per indicizzazione possono essere associati uno o più componenti di ricerca per indicizzazione.

DB dei collegamenti

Database dei collegamenti
In questo database vengono archiviate le informazioni estratte dal componente di elaborazione del contenuto e le informazioni sui clic.

DB di analisi

Database di report di analisi
Archivia i risultati dell'analisi dell'utilizzo.

Esempio di topologiadi ricerca

Farm multifunzione a tolleranza di errore per Ricerca contenuti organizzazione

Questa farm deve fornire un ambiente virtuale, completamente a tolleranza di errore per SharePoint Server 2016 inclusa la ricerca. Nell'illustrazione viene rappresentato l'esempio di una farm di ricerca contenuti organizzazione media con circa80 milioni di elementi nell'indice di ricerca.

Nota: *questo esempio non si applica a topologie di ricerca per siti Internet.*

Processi di analisi

L'architettura dell'analisi consiste nel **componente di elaborazione dei dati analitici**, **database di report dell'analisi** e **database dei collegamenti**.

3 Informazioni sul **componente di elaborazione dei dati analitici**

- Il componente di elaborazione dei dati analitici esegue due tipi di analisi, ovvero l'analisi della ricerca e l'analisi dell'utilizzo, e si basa sulle informazioni derivanti da queste analisi per migliorare la pertinenza della ricerca, creare report della ricerca e generare suggerimenti e collegamenti diretti.
- L'analisi della ricerca consiste nell'estrarre informazioni dal database dei collegamenti, ad esempio i collegamenti, il numero di volte che viene fatto clic su un elemento, il testo di ancoraggio, i dati correlati agli utenti e i metadati. Queste informazioni sono importanti per la pertinenza.
- L'analisi dell'utilizzo consiste nell'analizzare le informazioni dei log dei dati di utilizzo ricevute dal sistema front-end tramite l'archivio eventi. Quest'analisi genera report sull'uso e sulle statistiche.
- I risultati delle analisi vengono aggiunti agli elementi nell'indice di ricerca. I risultati dell'analisi dell'utilizzo inoltre vengono archiviati nel database di report di analisi.

B Informazioni sul **database dei collegamenti**

- Nel database dei collegamenti vengono archiviate informazioni estratte dal componente di elaborazione del contenuto. Vengono inoltre archiviate informazioni sui clic di ricerca e sul numero di clic eseguiti su un risultato di ricerca nella pagina dei risultati di ricerca. Le informazioni vengono memorizzate senza essere elaborate, per essere poi analizzate dal componente di elaborazione dei dati analitici.

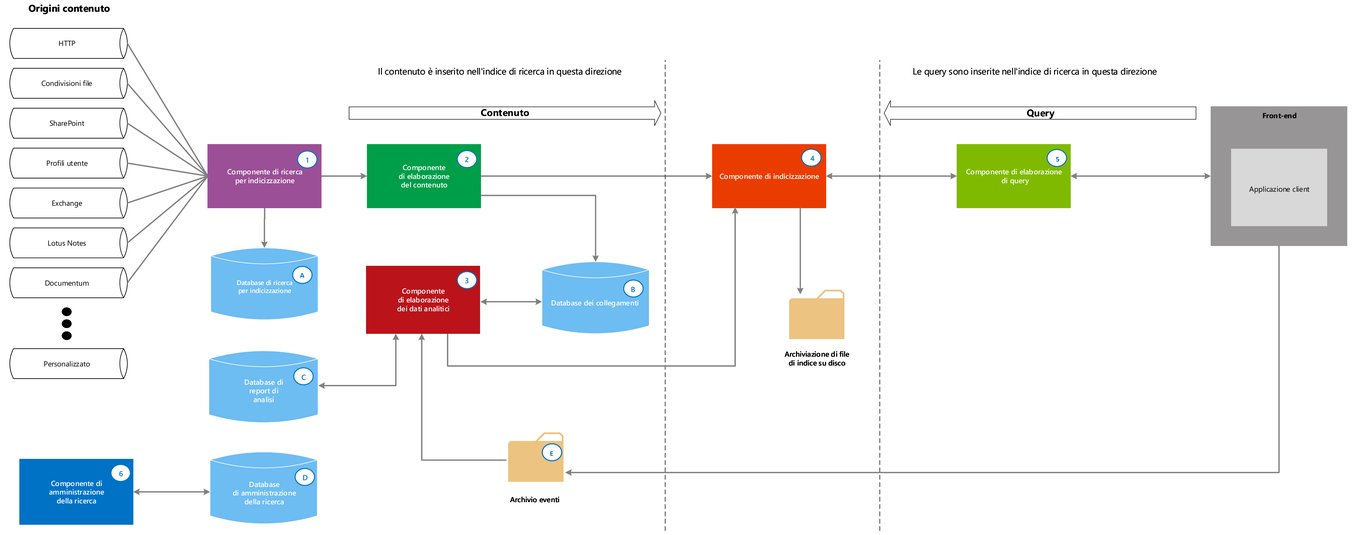
C Informazioni sul **database di report di analisi**

- Il database di report di analisi archivia i risultati dell'analisi dati di utilizzo.
- Inoltre, nel database di report di analisi vengono archiviate le informazioni statistiche dalle analisi. SharePoint usa queste informazioni per creare report Excel che mostrano statistiche diverse.

E Informazioni **sull'archivio eventi**

- L'archivio eventi contiene gli eventi di utilizzo che vengono acquisiti nel front-end, ad esempio il numero di volte in cui viene visualizzato un elemento.
- Questi eventi di uso vengono memorizzati come file di log sul server applicazioni che ospita il componente di elaborazione di analisi.

Illustrazione dei componenti di ricerca



SharePoint

© 2016 Microsoft Corporation. Tutti i diritti riservati. Per inviare commenti e suggerimenti su questa documentazione, scrivere all'indirizzo ITSPdocs@microsoft.com.

Microsoft