Search Architectures for SharePoint 2016

Overview

The search architecture in Microsoft SharePoint Server 2016 consists of components and databases that work together to perform the search operation. All components reside on application servers and all databases reside on database servers.

Search component interaction

Crawl and component processes

The crawl and component processes work together. Each component interacts with the crawl component, crawl database, and provisioning component. Both search components can be scaled out based on crawl volume and performance requirements.

About the crawl component

The crawl component is responsible for traversing content sources to deliver crawling threads to the actual search processes. Both the actual content and the associated metadata—the content processing component—must be correct.

About the crawl database

The crawl database contains detailed tracking and historical data about crawled items. This database holds information such as the crawl time, the last crawl ID, and the type of update during the last crawl.

About the content processing component

The content processing component is placed between the crawl component and the index component. It accepts crawling threads and sends these threads to the index component.

About the index component

An index component is the logical representation of an index replica. Each index replica is associated with a component name and a unique index ID.

About the index database

An index database is the logical representation of an index component. Each index replica is associated with a unique index ID and an index name.

Index and query processes

The index and query processes include the index component, index partition, and query component, all of which can be scaled out based on crawl volume, query volume, and performance requirements.

About the index component

An index component is the logical representation of an index replica in the search architecture. You have to provision one index component per each index replica.

About the index database

An index database is the logical representation of an index component. Each index database is associated with a unique index ID.

Application server with search components

An application server with search components consists of the application server, query processing component, and analytics processing component. It processes crawled items and feeds these items to the index component.

About the application server

Application servers are hosted on Windows servers. All SharePoint Databases are hosted on SQL Server.

Search components

Web server

Hosts Search Web Parts and Web Part pages for answering search queries in search farms. The application server is responsible for running the search administration component and for starting and stopping index partitions.

Search administration database

Stores search configuration data, such as the configuration, crawl rules, query rules, and mappings between content and managed properties.

Analytics processing component

The analytics processing component performs two types of analysis: search analytics and usage analytics. The search component runs information from these analyses to improve the search results.

Search analytics is the analysis of the documents on a SharePoint site and includes the following types of analyses:

- Keyword analysis
- Click analysis
- Query analysis
- Local site analysis

Usage analytics is the analysis of usage log information received from the front-end servers. This type of analysis includes the following types of analyses:

- Usage analysis
- Search analysis
- Query analysis
- Local site analysis

Example search topology

All purpose fault tolerant farm for Enterprise Search

This farm is designed to provide a fully fault tolerant, virtual environment for SharePoint Server 2016 including search. The illustration is an example of a virtual search farm with approximately 60 nodes in the search index. The hosts in this farm are eight CPU cores per virtual machine.

Note: This example also can apply to search topologies for Internet Sites.

Illustration of search component interaction

Web server

Hosts Search Web Parts and Web Part pages for answering search queries in search farms. The application server is responsible for running the search administration component and for starting and stopping index partitions. In small farms, this role can be shared on a server with the application server.

About the application server

Application servers are hosted on Windows servers. All SharePoint Databases are hosted on SQL Server.

Search administration database

Stores search configuration data, such as the configuration, crawl rules, query rules, and mappings between content and managed properties.

Analytics processing component

The analytics processing component performs two types of analysis: search analytics and usage analytics. The search component runs information from these analyses to improve the search results.

Search analytics is the analysis of the documents on a SharePoint site and includes the following types of analyses:

- Keyword analysis
- Click analysis
- Query analysis
- Local site analysis

Usage analytics is the analysis of usage log information received from the front-end servers. This type of analysis includes the following types of analyses:

- Usage analysis
- Search analysis
- Query analysis
- Local site analysis

Example search topology

All purpose fault tolerant farm for Enterprise Search

This farm is designed to provide a fully fault tolerant, virtual environment for SharePoint Server 2016 including search. The illustration is an example of a virtual search farm with approximately 60 nodes in the search index. The hosts in this farm are eight CPU cores per virtual machine.

Note: This example also can apply to search topologies for Internet Sites.

Illustration of search component interaction

Web server

Hosts Search Web Parts and Web Part pages for answering search queries in search farms. The application server is responsible for running the search administration component and for starting and stopping index partitions. In small farms, this role can be shared on a server with the application server.

About the application server

Application servers are hosted on Windows servers. All SharePoint Databases are hosted on SQL Server.

Search administration database

Stores search configuration data, such as the configuration, crawl rules, query rules, and mappings between content and managed properties.

Analytics processing component

The analytics processing component performs two types of analysis: search analytics and usage analytics. The search component runs information from these analyses to improve the search results.

Search analytics is the analysis of the documents on a SharePoint site and includes the following types of analyses:

- Keyword analysis
- Click analysis
- Query analysis
- Local site analysis

Usage analytics is the analysis of usage log information received from the front-end servers. This type of analysis includes the following types of analyses:

- Usage analysis
- Search analysis
- Query analysis
- Local site analysis

Example search topology

All purpose fault tolerant farm for Enterprise Search

This farm is designed to provide a fully fault tolerant, virtual environment for SharePoint Server 2016 including search. The illustration is an example of a virtual search farm with approximately 60 nodes in the search index. The hosts in this farm are eight CPU cores per virtual machine.

Note: This example also can apply to search topologies for Internet Sites.

Illustration of search component interaction

Web server

Hosts Search Web Parts and Web Part pages for answering search queries in search farms. The application server is responsible for running the search administration component and for starting and stopping index partitions. In small farms, this role can be shared on a server with the application server.

About the application server

Application servers are hosted on Windows servers. All SharePoint Databases are hosted on SQL Server.

Search administration database

Stores search configuration data, such as the configuration, crawl rules, query rules, and mappings between content and managed properties.

Analytics processing component

The analytics processing component performs two types of analysis: search analytics and usage analytics. The search component runs information from these analyses to improve the search results.

Search analytics is the analysis of the documents on a SharePoint site and includes the following types of analyses:

- Keyword analysis
- Click analysis
- Query analysis
- Local site analysis

Usage analytics is the analysis of usage log information received from the front-end servers. This type of analysis includes the following types of analyses:

- Usage analysis
- Search analysis
- Query analysis
- Local site analysis

Example search topology

All purpose fault tolerant farm for Enterprise Search

This farm is designed to provide a fully fault tolerant, virtual environment for SharePoint Server 2016 including search. The illustration is an example of a virtual search farm with approximately 60 nodes in the search index. The hosts in this farm are eight CPU cores per virtual machine.

Note: This example also can apply to search topologies for Internet Sites.