Enterprise Search Architectures
for SharePoint Server 2016 and SharePoint Server 2019

Overview

This model illustrates small, medium, and large-size farm architectures. The size of each farm is based on the number of items that are crawled and included in the search index. Architecture requirements can vary depending on the composition of the data that is crawled (size of items and formats). The examples illustrate the type of search components needed and how many of each. Use these examples as starting points for planning your own search environments. For more information about search processes and how search components interact, see Search Architectures for SharePoint® Server 2016 and SharePoint Server 2019 (https://go.microsoft.com/fwlink/p/?linkid=832554).

Search databases

Search admin DB

Search administration database
Stores search configuration data. Only one search administration database per Search service application.

Crawl DB

Crawl database
Stores the crawl history and manages crawl operations. Each crawl database can have one or more crawl components associated with it.

Link DB

Link database
Stores the information extracted by the content processing component and also stores click-through information.

Analytics DB

Analytics reporting database
Stores the results of usage analytics.

Search components

Index

Index component
The index component is the logical representation of an index replica.

Index partitions
- You can divide the index into discrete portions, each holding a separate part of the index.
- An index partition is stored in a set of files on a disk.
- The search index is the aggregation of all index partitions.

Index replicas
- Each index partition holds one or more index replicas that contain the same information.
- You have to provision one index component for each index replica.
- To achieve fault tolerance and redundancy, create additional index replicas for each index partition and distribute the index replicas over multiple application servers.

Query processing

Query processing component
Analyzes and processes search queries and results.

Admin

Search administration component
Runs system processes that are essential to search. There can be more than one search administration component per Search service application, but only one is active at any given time.

Crawl

Crawl component
Crawls content based on what is specified in the crawl databases.

Content processing

Content processing component
Carries out various processes on the crawled items, such as document parsing and property mapping.

Analytics

Analytics processing component
Carries out search analytics and usage analytics.

Hardware requirements and scaling considerations

The requirements apply to each of the servers in the small, medium or large Enterprise Search topology. You can deploy search topologies for the enterprise on physical hardware or on virtual machines.

Note: *For evaluation use, you can place all search components on one server with 8GB RAM.*

Minimum hardware requirements for application servers

SEARCH COMPONENT	HARD DISK	RAM	PROCESSOR
Index component	80GB regardless of the number of search components hosted on the server.* 500GB additional disk space, preferably a separate disk volume/partition.	32GB if the server hosts only an index component. 32GB if the server hosts an index component and a query processing component.	64-bit, 8 cores minimum. When hosting virtual machines on Windows Server 2008 R2 SP1, maximum 4 cores are possible.
Analytics processing component	80GB regardless of the number of search components hosted on the server.* 300GB additional disk space, preferably a separate disk volume/partition. This disk space is necessary for local processing of analytics data before it is written to the analytics reporting database.	8GB if the server hosts only one of these search components. 16GB if the server hosts two or more of these search components. This does not apply if the server hosts an index component.	64-bit, 4 cores minimum, 8 cores recommended. When hosting virtual machines on Windows Server 2008 R2 SP1, maximum 4 cores are possible.
Crawl component Content processing component Query processing component Search administration component	80GB regardless of the number of search components hosted on the server.*		

* The server must have sufficient disk space for the base installation of the Windows Server operating system and sufficient disk space for diagnostics such as logging, debugging, creating memory dumps, and so on.

For production use, the server also needs additional free disk space for day-to-day operations and for the page file. Follow the guidance on free disk space and page file size corresponding to your Windows Server installation.

Minimum hardware requirements for database servers

HARDWARE COMPONENT	REQUIREMENTS
Processor	64-bit, 4 cores for small deployments. 64-bit, 8 cores for medium deployments.
RAM	8 GB for small deployments. 16 GB for medium deployments.
Hard Disk	80 GB for system drive.*

Scaling out for performance

Key performance metrics and scale actions

TO IMPROVE THIS METRIC	TAKE THESE ACTIONS
Full crawl time and result freshness	Add more crawl databases and content processing components for result freshness. Crawl databases and content processing components can be distributed among their respective servers. The Crawl Health Reports can be used to determine the cause of bottlenecks, if any.
Time required for results to be returned	**To improve query latency:** Add more index replicas so that the query load is distributed more evenly among the index replicas. This solution is better suited for small topologies. **To improve query latency and query throughput:** Split the search index into more partitions to reduce the number of items on each partition. The Query Health Reports can be used to determine the cause of bottlenecks, if any.
Availability of query functionality	Deploy redundant (failover) query processing components on different application servers.
Availability of content crawling, processing, and indexing functionality	Use multiple crawl databases on redundant database servers. Use multiple content processing components on redundant application servers.

Scaling out search components as number of items increase

NUMBER OF ITEMS	INDEX COMPONENTS AND PARTITIONS	QUERY PROCESSING COMPONENTS	CONTENT PROCESSING COMPONENTS	ANALYTICS PROCESSING COMPONENTS	CRAWL COMPONENTS	CRAWL DATABASES	LINK DATABASE	ANALYSTICS REPORTING DATABASE	SEARCH ADMINISTRATION COMPONENT
General Guidance	Add 1 index partition per 20 million items	Use 2 query processing components for redundancy. Above 80 million items, increase to 4.				Add 1 crawl database per 20 million items	Add 1 link database per 60 million items	Add an analytics reporting database when analytics performance starts decreasing	Use 2 search administration components for redundancy, for all farm sizes
20 million	2 components 1 partition	2	2	2	2	1	1	Variable	2
20-80 million	8 components 4 partitions	2	4	2	2	2	1	Variable	2
200 million	20 components 10 partition	4	6	6	2	5	2	Variable	2

Redundancy and availability

Redundant search components must be installed on separate failure domains. All of the example topologies: Small, Medium, and Large have redundant configurations.

Search database redundancy must be handled by the SQL server configuration. SQL 2008 R2 and SQL 2012 are supported.

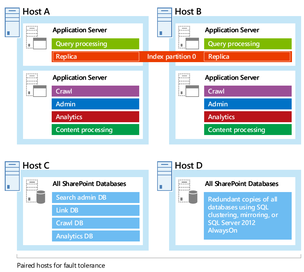
For redundant crawling and query processing, it is not necessary to have a redundant analytics processing component. However, if the non-redundant analytics processing component fails, the search results will not have optimal relevance until the failure is recovered.

Example search topologies

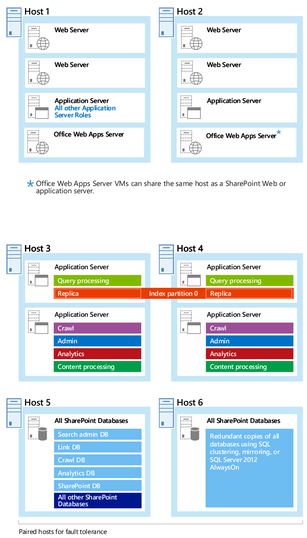
Small search farm (~20M Items)

This farm is intended to provide the full functionality of SharePoint Server search with fault tolerance for up to 20 million items in the search index. Two versions are illustrated..

Dedicated search farm
This farm illustrates only the search components and can serve as a dedicated search farm for one or more SharePoint farms. Dedicated search farms do not include Web servers.



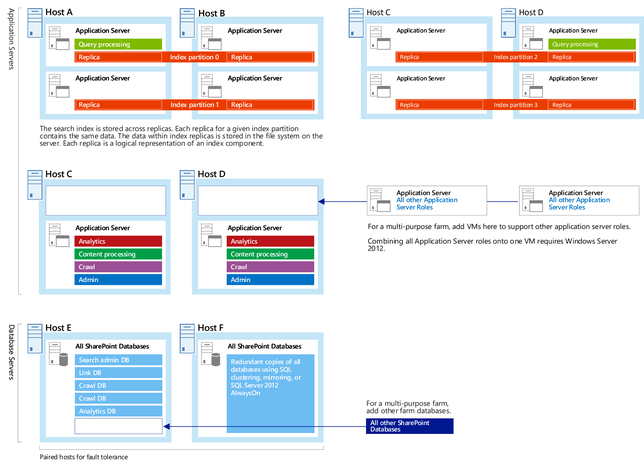
All purpose farm
This farm incorporates the full functionality of SharePoint Server.



* Office Web Apps Server VMs can share the same host as a SharePoint Web or application server.

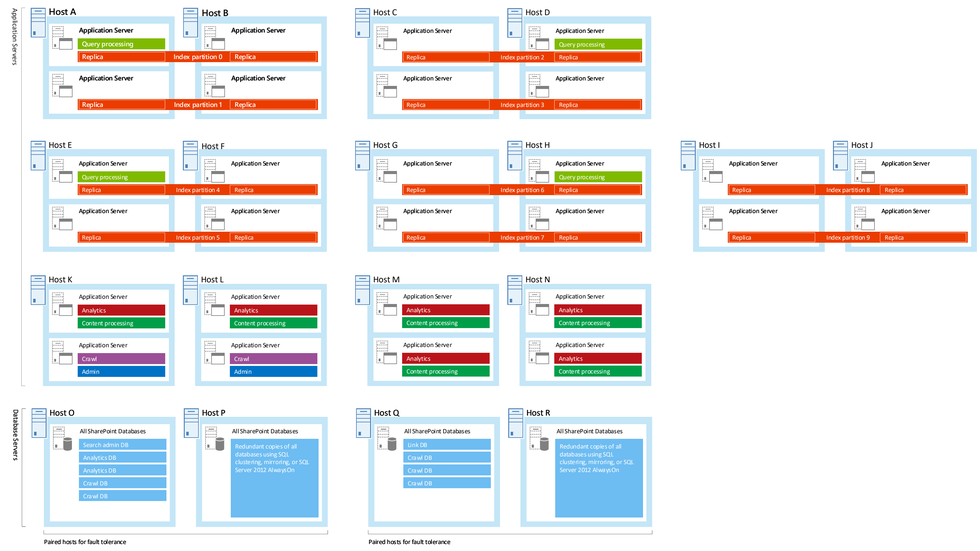
Medium search farm (~80M Items)

This farm is intended to provide the full functionality of SharePoint Server search with fault tolerance for up to 80 million items in the search index. To make this an all-purpose farm, add Web servers (not shown) and the additional application servers and databases that are noted.



Large search farm (~200M Items)

This farm is intended to provide a dedicated search farm with fault tolerance for up to 200 million items in the search index.



SharePoint

Microsoft