

Архитектуры поиска на веб-сайтах

для SharePoint Server 2016

Общие сведения

Эта модель иллюстрирует архитектуру среднего размера для фермы поиска на веб-сайтах. Требования к архитектуре зависят от состава данных (размеров и форматов элементов), обход которых выполняется. В этом примере показаны типы необходимых компонентов поиска и их количество. Используйте этот пример в качестве отправной точки для планирования собственных сред поиска. Дополнительные сведения о процессах и взаимодействии компонентов поиска см. на афише "Архитектуры поиска для SharePoint® Server 2016" (<https://go.microsoft.com/fwlink/p/?linkid=832554>).

Компоненты поиска

Индекс
Компонент индексирования Компонент индексирования — это логическое представление реплики индекса.
Разделы индекса <ul style="list-style-type: none">Индекс можно разделить на фрагменты, каждый из которых отвечает за его отдельную часть.Раздел индекса хранится в виде набора файлов на диске.Индекс поиска — это совокупность всех разделов индекса.
Реплики индекса <ul style="list-style-type: none">Каждый раздел индекса содержит одну или несколько реплик с одинаковыми данными.Для каждой реплики индекса необходимо подготовить по одному компоненту индексирования.Чтобы обеспечить отказоустойчивость и избыточность, создайте дополнительные реплики индекса для каждого его раздела и распределите эти реплики индекса между несколькими серверами приложений.
Обработка запросов
Компонент обработки запросов Анализирует и обрабатывает поисковые запросы и результаты.
Администратор
Компонент администрирования поиска Выполняет системные процессы, необходимые для поиска. В приложении службы поиска может быть несколько компонентов администрирования поиска, но они не могут быть активными одновременно.
Обход контента
Компонент обхода Выполняет обход контента на основании сведений, указанных в базах данных обхода.
Обработка контента
Компонент обработки контента Выполняет различные операции с элементами, для которых выполнен обход (например, разбор документа и сопоставление свойств).
Аналитика
Компонент обработки аналитики Выполняет анализ поиска и использования.

Базы данных поиска

администрирования
База данных администрирования поиска Хранит данные конфигурации поиска. У каждого приложения службы поиска может быть только одна база данных администрирования поиска.
БД обхода
База данных обхода Хранит журнал обхода контента и управляет операциями обхода. С каждой базой данных обхода может быть связан один или несколько компонентов обхода.
БД ссылок
База данных ссылок Хранит данные, извлеченные компонентом обработки контента, а также сведения о выборе рекламных объявлений.
БД аналитики
База данных отчетов аналитики Хранит результаты аналитики по использованию.

Рекомендации по производительности

Серверы приложений, на которых размещаются компоненты поиска

ВАЖНЫЕ ФАКТОРЫ	ОБЪЯСНЕНИЕ
Кэш	Запрос и его результаты кэшируются Windows Server AppFabric парами ключ-значение. 1 де кило — это запрос, а значение — результаты. Для каждого запроса коэффициент кэширования составляет примерно 50 %. Это значит, что при обработке 200 запросов в секунду примерно 100 из них отправляются в индекс поиска, а остальные — кэшируются. Для результатов из кэша задержка обработки запросов меньше, чем для результатов из индекса поиска. Например, результаты повторяющихся запросов по титульной странице обычно кэшируются.
Непрерывный обход контента	Рекомендуем включить непрерывный обход контента с интервалом в 1 минуту (по умолчанию он равняется 15 минутам). Активировать его можно только для источников контента SharePoint.
Анонимный доступ	При анонимном входе пользователям не нужно указывать учетные данные для входа на веб-сайт SharePoint. Кроме того, анонимные запросы кэшируются, поэтому они менее затратны благодаря меньшей задержке их обработки. Анонимный доступ следует включить как в интерфейсном веб-компоненте, так и на самом сайте.
Задержка обработки запросов	На задержку обработки запросов влияет кэширование, анонимный доступ, количество и сложность применяемых правил и другие факторы. Среди них — диски, на которых хранится индекс поиска. Если у них несколько шинделей, то скорость доступа будет выше, а задержка обработки запросов — меньше.

Топологии поиска для веб-сайтов

Нижe приведены отличия топологий для веб-сайтов от других топологий поиска.

- Веб-серверы внедряются в топологию веб-сайтов для обслуживания запросов пользователей. В этом она отличается от топологии поиска в корпоративной среде, где веб-серверы не входят в состав выделенной фермы поиска.
- Компонент обработки запросов и компоненты индексирования размещаются на веб-серверах, обеспечивая максимально эффективное использование имеющихся аппаратных ресурсов и упрощая масштабирование топологии поиска.

Требования к оборудованию

Эти требования относятся к топологии среднего размера для поиска на веб-сайтах. Минимальные требования к ОЗУ для сервера, на котором размещается компонент поиска, — это общий объем ОЗУ, необходимый для этого сервера. Например, при размещении компонента обработки контента, компонента администрирования поиска и компонента обхода на одном сервере общий объем ОЗУ, необходимый для этого сервера, составляет 24 Гб.

Минимальные требования к оборудованию для серверов приложений

КОМПОНЕНТ ПОИСКА НА ФИЗИЧЕСКОМ СЕРВЕРЕ	ОЗУ	ЖЕСТКИЙ ДИСК	ПРОЦЕССОР
Компонент индексирования	48 Гб для каждого сервера в ферме, где размещается компонент индексирования, компонент обработки запросов и	500 Гб дополнительного места на диске, желательно на отдельном томе или разделе диска.	ВСЕ КОМПОНЕНТЫ: 64-разрядный, не менее 4 ядер, рекомендуется 8.
Компонент обработки аналитики	24 Гб для каждого сервера фермы, на котором размещены компоненты обработки аналитики, обхода и/или администрирования поиска.	300 Гб дополнительного места на диске, желательно на отдельном томе или разделе диска. Это место на диске необходимо для локальной обработки данных аналитики до записи в базу данных отчетов.	
Компонент обхода		80 Гб для системного диска.	
Компонент обработки контента	См. требования для компонента обработки аналитики.	У вас должно быть достаточно места для основной установки и диагностики, например ведения журнала, отладки, создания дампов памяти и других операций. Для рабочего использования также требуется дополнительное свободное пространство для повседневных операций. В рабочих средах постоянно поддерживайте объем свободного пространства, в два раза превышающий объем ОЗУ.	
Компонент обработки запросов	См. требования, указанные для компонента индексирования.		
Компонент администрирования поиска	См. требования для компонента обработки аналитики.		

Минимальные требования к оборудованию для серверов баз данных

КОМПОНЕНТ ОБОРУДОВАНИЯ	ТРЕБОВАНИЯ
Процессор	64-разрядный, 4 ядра для небольших развертываний 8 Гб для небольших развертываний
ОЗУ	16 Гб для средних развертываний
Жесткий диск	80 Гб на системном диске

Ферма среднего размера для веб-сайтов (около 85 просмотров страниц в секунду)

Эта ферма предназначена для создания отказоустойчивой топологии фермы поиска SharePoint Server 2016, оптимизированной для собрания из 3 400 000 элементов. Пример фермы обрабатывает 100–200 документов в секунду (в зависимости от языка) и обеспечивает 85 просмотров страниц в секунду и 100 запросов в секунду.

Эта топология оптимизирована для физического оборудования, но ее можно развернуть и на виртуальных машинах.

