

Internet 站点搜索体系结构

适用于 SharePoint Server 2016

概述

本模型展示了中型 Internet 站点体系结构。体系结构要求可能因已爬网的数据组成（项大小和格式）而异。此示例展示了所需的搜索组件类型，以及每个类型所需的数量。请从此示例入手，规划你自己的搜索环境。若要详细了解搜索处理以及搜索组件的交互方式，请参阅“SharePoint® Server 2016 的搜索体系结构” (<https://go.microsoft.com/fwlink/p/?linkid=832554>)。

搜索组件

索引

索引组件
索引组件是索引副本的逻辑表示。

索引分区

- 可以将索引划分为各个独立部分，每个保留索引的不同部分。
- 索引分区存储在磁盘上的一组文件中。
- 搜索索引汇总了所有索引分区。

索引副本

- 每个索引分区保留一个或多个包含相同信息的索引副本。
- 必须为每个索引副本预配一个索引组件。
- 为了实现容错和冗余，请为每个索引分区创建其他索引副本，并在多个应用程序服务器上分布索引副本。

查询处理

查询处理组件
分析并处理搜索查询和结果。

管理

搜索管理组件
执行搜索所必需的系统处理。每个 Search Service 应用程序可以有多个搜索管理组件，但在任何指定时间，只有一个可用。

爬网

爬网组件
根据爬网数据库中指定的内容对内容进行爬网。

内容处理

内容处理组件
对已爬网项执行各种处理（如文档分析和属性映射）。

分析

分析处理组件
执行搜索分析和使用情况分析。

搜索数据库

搜索管理 DB

搜索管理数据库
存储搜索配置数据。每个 Search Service 应用程序只有一个搜索管理数据库。

爬网 DB

爬网数据库
存储爬网历史记录并管理爬网操作。每个爬网数据库可以有一个或多个与之相关联的爬网组件。

链接 DB

链接数据库
存储由内容处理组件提取的信息，也存储单击后到达信息。

分析 DB

分析报表数据库
存储使用情况分析结果。

性能注意事项

托管搜索组件的应用程序服务器

注意事项	原因
缓存	查询及其结果通过 Windows Server AppFabric 缓存存储在键值对中：查询是键，而结果是值。对于每个查询，缓存率约为 50%。这意味着，如果使用模式为每秒 200 个查询，则会将大约 100 个查询发送到搜索索引，而其余 100 个查询将进行缓存。与搜索索引中的结果相比，缓存中的结果具有较低的查询延迟。例如，重复运行的标题页查询的结果可能要进行缓存。
连续爬网	我们建议按 1 分钟的时间间隔启用连续爬网，而不是默认的 15 分钟时间间隔。只能对 SharePoint 内容源启用连续爬网。
匿名访问	通过匿名访问，用户无需使用凭据即可登录 SharePoint Internet 站点。此外，将对匿名查询进行缓存，因此，由于查询延迟较低，匿名查询更为划算。必须在两个位置启用匿名访问：Web 前端和站点本身。
查询延迟	查询延迟受缓存、匿名访问以及应用和触发的查询规则的数量和复杂性等其他因素的影响。此外，请考虑存储搜索索引的磁盘；具有多个主轴的磁盘可以提高磁盘的访问速度并降低查询延迟。

Internet 站点的搜索拓扑

Internet 站点的搜索拓扑不同于其他搜索拓扑，区别如下：

- Web 服务器包含在 Internet 站点的搜索拓扑中，以便处理用户请求。这与企业级搜索的搜索拓扑不同。在后者的搜索拓扑中，Web 服务器并不属于专用搜索服务器场。
- 查询处理组件和索引组件驻留在 Web 服务器上，以最大限度地利用可用硬件资源，并简化搜索拓扑扩展。

硬件要求

这些要求适用于中型 Internet 站点搜索拓扑。对于托管搜索组件的服务器，列出的最低 RAM 要求是该服务器一共需要的 RAM 量。例如，如果你要在同一服务器上托管内容处理组件、搜索管理组件和爬网组件，该服务器一共至少需要 24 GB RAM。

应用程序服务器的最低硬件要求

物理服务器上的搜索组件	RAM	硬盘	处理器
索引组件	48 GB，用于服务器场中托管索引组件、查询处理组件和 Web 前端的每台服务器。	500 GB 额外磁盘空间，最好是单独的磁盘卷/分区。	所有组件： 64 位，最少 4 核，建议为 8 核。
分析处理组件	24 GB，用于场中托管分析处理组件、爬网组件和/或搜索管理组件的每台服务器。	300 GB 额外磁盘空间，最好是单独的磁盘卷/分区。在将分析数据写入分析报表数据库之前，必须有此类额外的磁盘空间，才能在本地处理分析数据。 80GB，用于系统驱动器。 磁盘空间必须足以容纳基础安装以及诊断（如日志记录和测试）、创建内存转储等其他操作。若要满足生产用途，还需要为日常操作提供额外的可用磁盘空间。可用空间应维持在用于生产环境的 RAM 的两倍。	
爬网组件内容处理组件	请参阅针对分析处理组件列出的要求。		
查询处理组件	请参阅针对索引组件列出的要求。		
搜索管理组件	请参阅针对分析处理组件列出的要求。		

数据库服务器的最低硬件要求

硬件组件	要求
处理器	64 位，4 个内核（用于小型部署） 8 GB，用于小型部署
RAM	16 GB，用于中型部署
Hard Disk	80GB for system drive

中型 Internet 站点场（每秒约 85 个页面视图）

此场旨在提供容错 SharePoint Server 2016 搜索服务器场拓扑，该拓扑已针对包含 3,400,000 项的语料库进行了优化。示例场每秒可处理 100-200 个文档（具体视语言而定），每秒可容纳 85 个页面视图和 100 个查询。

虽然该拓扑已针对物理硬件进行了优化，但你也可以在虚拟机上进行部署。

