

企业级搜索体系结构

适用于 SharePoint Server 2016

概述

本模型展示了小型、中型和大型体系结构。每个场的大小均以搜索索引中已爬网和包含的项数为依据。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。本文中的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。请从这些示例入手，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。

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，但在任何给定时间，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。

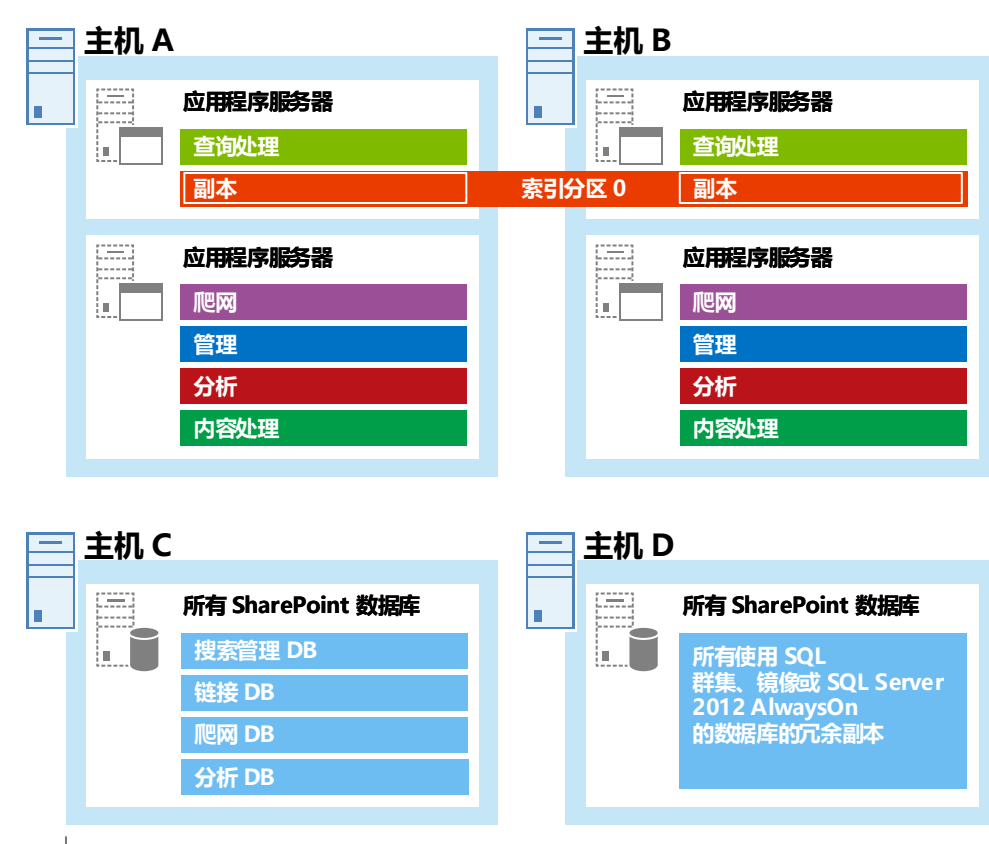
搜索拓扑示例

小型搜索服务器场（大约 2000 万个项）

此场旨在提供 SharePoint Server 2016 搜索的全部功能，为搜索索引中多达 2000 万个项提供容错。展示了两个版本。

专用的搜索服务器场

此场仅展示了搜索组件，可作为一个或多个 SharePoint 场的专用搜索服务器场。专用搜索服务器场不包括 Web 服务器。



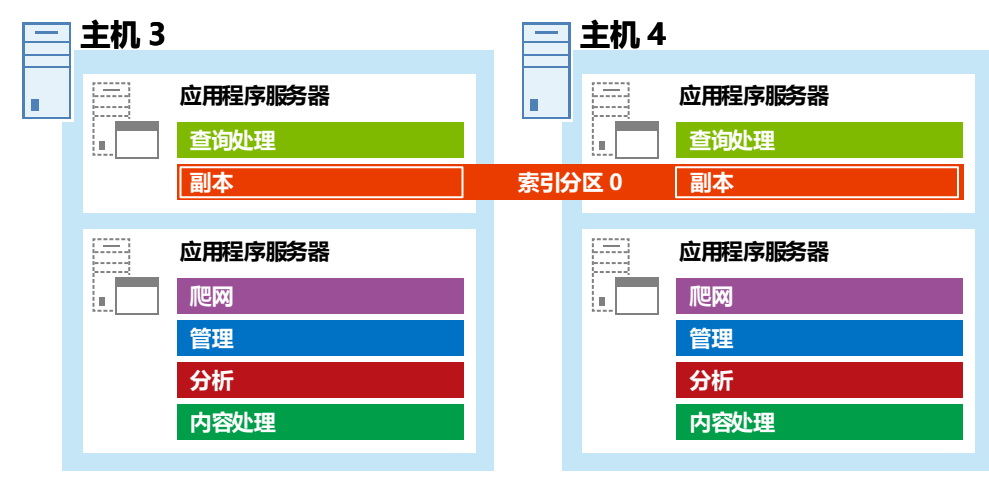
为实现容错的配置主机

多用途场

此场包含 SharePoint Server 2016 的全部功能。



* Office Web Apps Server VM 可以与 SharePoint Web 或应用程序服务器共用主机。



为实现容错的配置主机

中型搜索服务器场（大约 8000 万个项）

此场旨在提供 SharePoint Server 2016 搜索的全部功能，为搜索索引中多达 8000 万个项提供容错。若要使其成为多用途场，请添加 Web 服务器（未显示）以及注明的其他应用程序服务器和数据库。



搜索索引副本存储。指定索引分区的每个副本均包含完整数据。索引副本内的数据存储在服务器上的文件夹中。每个副本都是索引组件的逻辑表示。

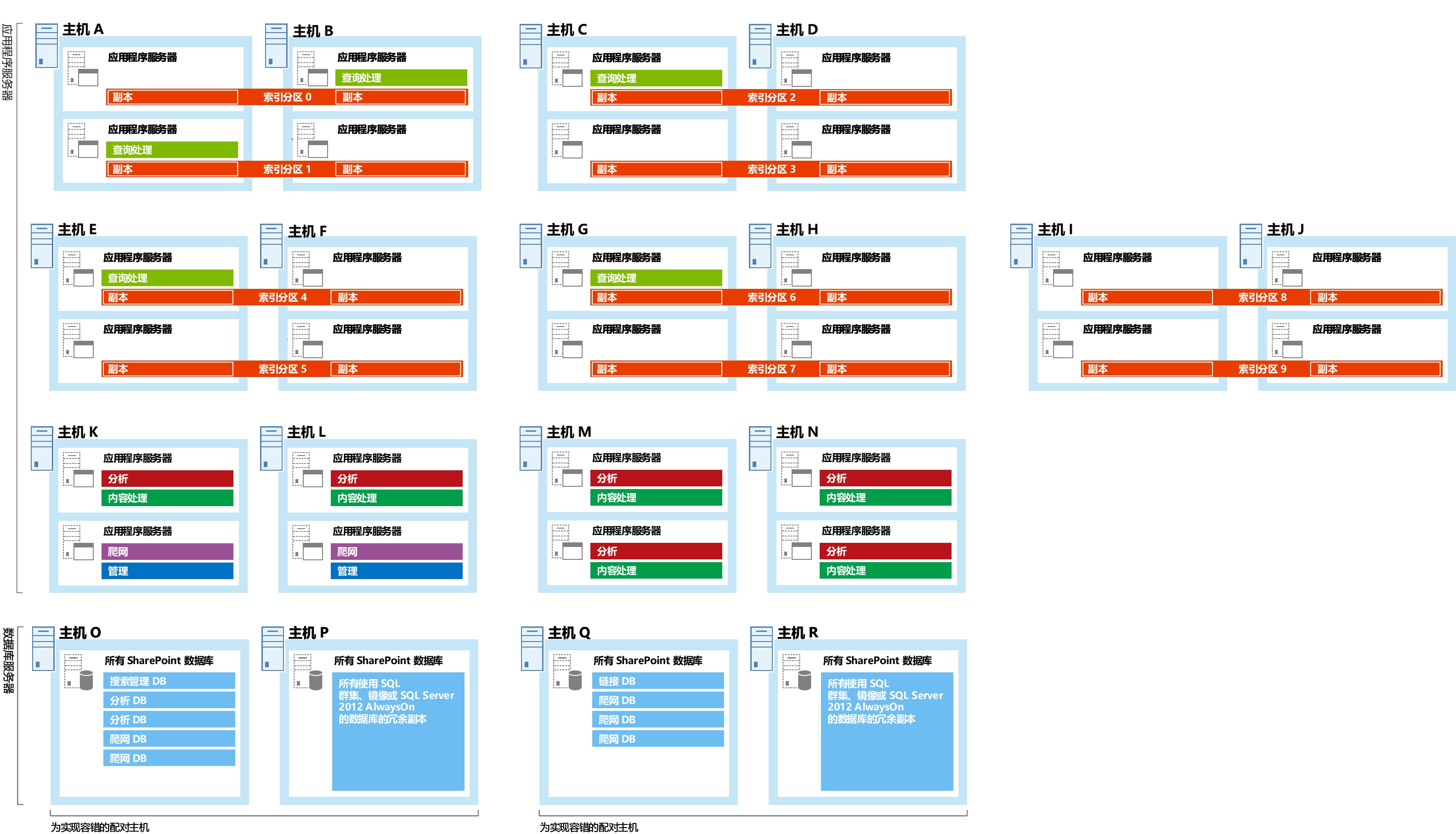
对于多用途场，在此处添加 VM，支持其他应用程序服务器角色。必须安装 Windows Server 2012，才能将所有应用程序服务器角色合并到一个 VM 中。

对于多用途场，添加其他数据库。其他所有 SharePoint 数据库。

为实现容错的配置主机

大型搜索服务器场（大约 2 亿个项）

此场旨在提供专用的搜索服务器场，为搜索索引中多达 2 亿个项提供容错。



为实现容错的配置主机

数据库服务器的最低硬件要求

硬件组件	要求
处理器	64 位，对于小型部署，为 4 核。 64 位，对于中型部署，为 8 核。 8 GB，用于小型部署。
RAM	16 GB，用于中型部署。
硬盘	80 GB，用于系统驱动器。*

扩展以提升性能

关键性能指标和扩展操作

若要提高此指标	请采取以下措施
完全爬网时间和结果新鲜度	请采取以下措施： 添加更多爬网数据库和内容处理组件以保持结果新鲜度。 爬网数据库和内容处理组件可以分布在其各自的服务器上。 可以使用爬网运行状况报告确定导致瓶颈（如果有）产生的原因。
返回结果所需的时间	若要改善查询延迟并提高查询吞吐量，请采取以下措施： 将搜索索引分成更多分区，以减少每个分区上的项数。 可以使用查询运行状况报告确定导致瓶颈（如果有）产生的原因。
查询功能的可用性	在不同的应用程序服务器上部署冗余（故障转移）查询处理组件。 在冗余数据库服务器上使用多个爬网数据库。
内容爬网、处理和索引功能的可用性	在冗余应用程序服务器上使用多个内容处理组件。

冗余和可用性

冗余搜索组件必须安装在单独的故障域上。所有示例拓扑（小型、中型和大型）都有冗余配置。

搜索数据库冗余必须由 SQL 服务器配置处理。支持 SQL 2008 R2 和 SQL 2012。

对于冗余爬网和查询处理，无需使用冗余分析处理组件。不过，如果非冗余分析处理组件失败，那么在故障恢复前搜索结果的相关性会不佳。