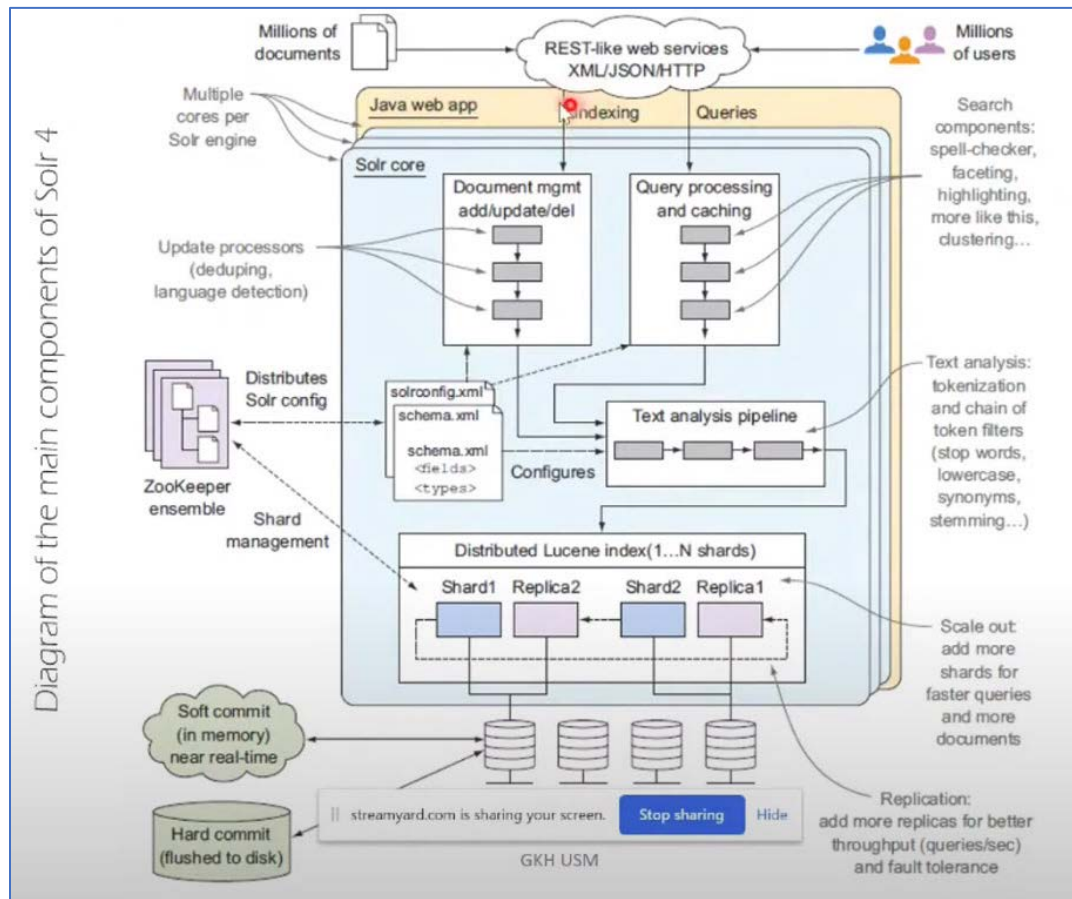


Content

1.	Solr and ZooKeeper Relations	--- 2
2.	Solr system Requirements	--- 3
3.	Sitecore 9.3 system Requirements – XC	--- 4
	• Hardware Requirements	
	• Hosting Environment Requirements	
4.	Sitecore HA Backend Deployment	--- 6
5.	Public Indexing Service - Solr Performance Data	--- 7
6.	Java Platform - Runtime	--- 8
7.	SolrCloud Model - Example	--- 9
8.	Pros & Cons for Solr	--- 9
9.	Solr vs. Solrcloud	--- 10
10.	Solrcloud Questionnaire	--- 11

SOLR AND ZOOKEEPER RELATIONS



UI application

Hue

Zoom data

Banana

Blacklight

Velocity UI

AJAX Solr

Search platform service

Solr

Logic enhancement

API/Handler/Analysis

Rule engine

Carrot2

Indexing & searching

Lucene

Document building

API/
Handler
/Analysis
s

Morphlines

Lily

Solandra

Tika

UIMA

OpenNLP

GATE

Nutch

ManifoldCF

Spatial4J

Data

*

Flume

NoSQL
Hbase

Cassandra

Document
/File

Natural Language

Web site

Content
repository

Spatial

SOLR SYSTEM REQUIREMENTS

https://solr.apache.org/guide/8_4/solr-system-requirements.html

<https://cwiki.apache.org/confluence/display/solr/solrperformanceproblems>

HARDWARE REQUIREMENTS

Solr runs on a 64-bit Java.

A 64-bit Java requires a 64-bit operating system.

SUPPORTED OPERATING SYSTEMS

Solr is tested on several versions of Linux, macOS and Windows.

JAVA REQUIREMENTS

You will need **the Java Runtime Environment (JRE) version 1.8 or higher.**

```
$ java -version
java version "1.8.0_60"
Java(TM) SE Runtime Environment (build 1.8.0_60-b27)
Java HotSpot(TM) 64-Bit Server VM (build 25.60-b23, mixed mode)
```

LUCENE/SOLR 8.X

- Requires **Java 8** or higher.
- This version has continuous testing with Java 9, 10, 11, 12 and the pre-release version of Java 13.
- There were known issues with Kerberos with Java 9+ prior to Solr 8.1. If using 8.0, you should test in your environment.
- Be sure to test with SSL/TLS and/or authorization enabled in your environment if you require either when using Java 9+.

SITECORE 9.3 SYSTEM REQUIREMENTS - XC

<https://doc.sitecore.com/en/developers/93/sitecore-experience-commerce/system-requirements.html>

(Sitecore Version 10 does not have system requirements information)

HARDWARE REQUIREMENTS

- 4 core processor
- 16 GB of RAM

HOSTING ENVIRONMENT REQUIREMENTS

The following table lists the software requirements for Sitecore XC 9.3 hosting environment.

<i>Operating system</i>	<ul style="list-style-type: none">• Windows Server 2016• Windows Server 2019• Windows 10 Professional (64-bit)
<i>.NET Framework</i>	<ul style="list-style-type: none">• .NET Framework 4.7.1 Developer Pack (including available updates)• ASP.NET MVC 5.2 or later• ASP.NET Core 2.1• .NET Core Windows Server Hosting 2.1.6 or later versions of 2.1• OData 7.2.0
<i>Database</i>	<ul style="list-style-type: none">• Microsoft SQL Server 2019• Microsoft SQL Server 2017• Microsoft SQL Server 2016 SP2• Microsoft Azure SQL• Redis data store• Note Sitecore recommends using Redis for Linux for production installations. The version for Windows should only be used for development or non-production single machine deployments.
<i>Web server</i>	<ul style="list-style-type: none">• IIS 10.0 FC: (XP0 9.3)• Microsoft Web Deploy 3.6

SolrCloud Questionnaire

	<ul style="list-style-type: none">• URL Rewriter
<i>Development</i>	<ul style="list-style-type: none">• Visual Studio 2017 + latest updates• Visual Studio 2017 Tooling• Visual Studio 2019• Visual Studio 2019 Tooling• .NET Core 2.1.7 Windows Hosting module• MSBuild Microsoft Visual Studio Web targets (available from Nuget)• Git client
<i>Deployment</i>	PowerShell 6.0 or later
<i>Runtime</i>	<ul style="list-style-type: none">• .Net Core Runtime 2.1• Node.js• OData Connected Service (available from Visual Studio Marketplace)
<i>Search indexing</i>	Solr 8.1.1 → 8.4.0 for Sitecore 10.1
<i>Sitecore software</i>	<ul style="list-style-type: none">• Sitecore Experience Platform 9.3 (available on dev.sitecore.net)• Sitecore Experience Accelerator (SXA) 9.3 (available on dev.sitecore.net)• Sitecore PowerShell Extensions 6.0 for Sitecore 9.3 (available with SXA downloads on dev.sitecore.net)• Sitecore Identity 4.0.0• Note Sitecore SXA and Sitecore PowerShell Extensions are installed when you run the Sitecore Experience Commerce deployment script, as part of the XC installation process.
<i>Web browsers</i>	<ul style="list-style-type: none">• Microsoft Edge• Internet Explorer 11• Mozilla Firefox• Google Chrome• Apple Safari

SITECORE HA BACKEND DEPLOYMENT

Reference : Professional Sitecore 8 Development.pdf, Phil Wicklund & Jason Wilkerson

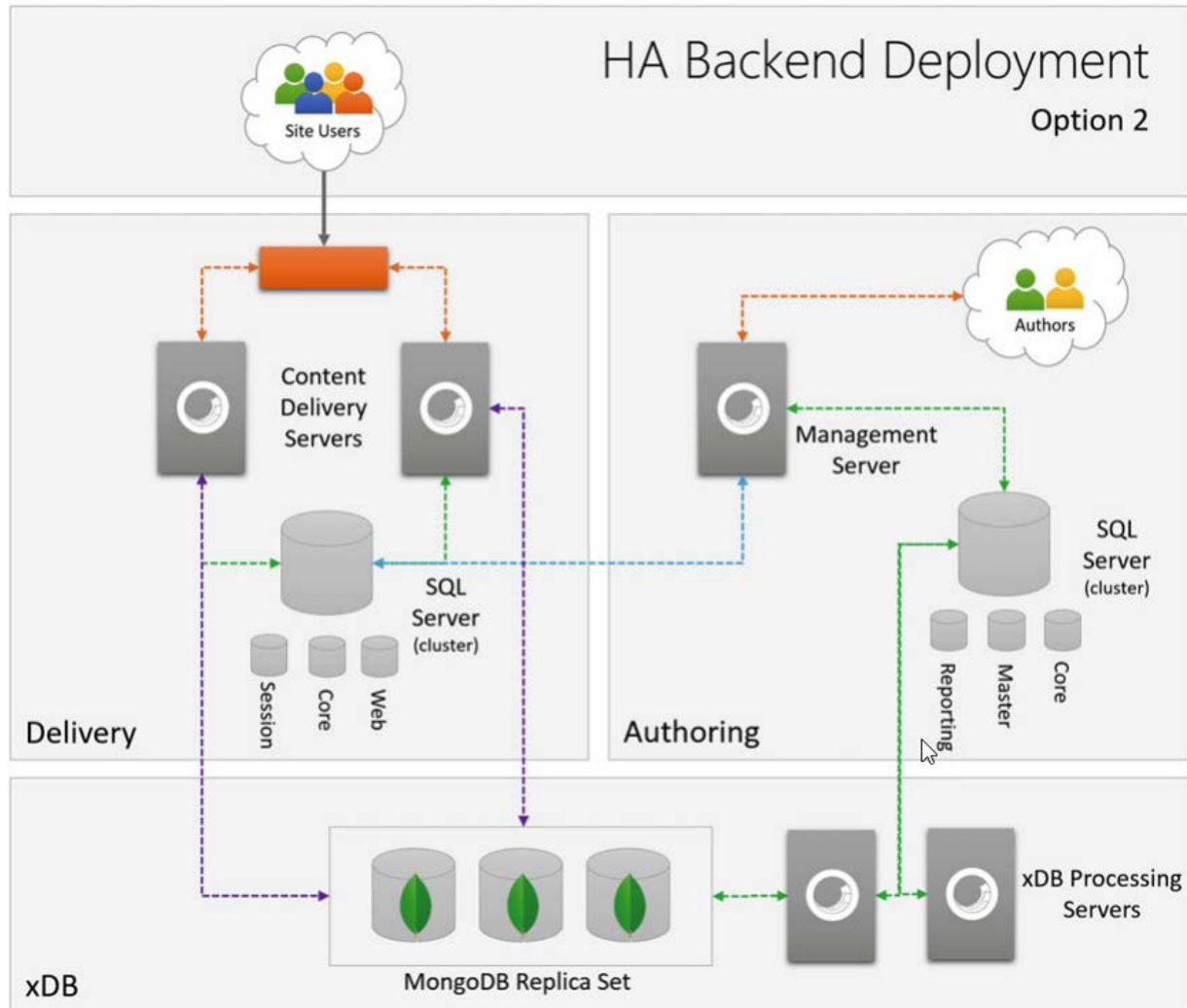


Figure 1-6. Option 2 features high availability through the solution, save the management server

SolrCloud Questionnaire

SOLR PERFORMANCE DATA

Created by ASF Infrabot on Jun 28, 2019

<https://cwiki.apache.org/confluence/display/SOLR/SolrPerformanceData>

CNET SHOPPER.COM - STANDALONE

- **2.4GHz dual opteron (DL385) w/ 16GB memory Linux (2.6.9).**
- index used in these tests contained ~400K records, took up ~900MB of disk
- a cron job forcibly triggered a commit (even though the index hadn't changed) every 15 minutes to force a new searcher to be opened and autowarmed while the queries were being processed

DISCOGS.COM - STANDALONE

- **2.66Ghz P4, with 1 gig ram**
- index has about 1.2 million documents, 1.2 gigs in size.
- This machine handles 250,000 queries per day

HATHITRUST - SOLRCLOUD

- Response time is around 200 ms
- 5 million full-text books indexed
- **production index is spread across 10 shards on 4 machines**
- total index size of over 2 Terabytes

ZVENTS - SOLRCLOUD

- 8 millions users monthly with engaging local contents
- **queries are served by a cluster of 8 machines, each having 16Gigs of memory and 4 cores**
- Our search index contains over 4 million documents
- total index size of over 2 Terabytes

DANISH WEB ARCHIVE - SINGLE SOLRCLOUD

- Archive with 500TB+ harvested web resources
- A single 24 core 256GB CentOS machine builds the shards
- Solr indexer has 32GB heap
- **All instances are in a single SolrCloud**

SolrCloud Questionnaire

JAVA PLATFORM - RUNTIME

<https://java.com/en/download/help/sysreq.html>

<https://www.oracle.com/java/technologies/javase/products-doc-jdk8-jre8-certconfig.html>

WINDOWS

- Windows 10 (8u51 and above)
- Windows 8.x (Desktop)
- Windows 7 SP1
- Windows Vista SP2
- Windows Server 2008 R2 SP1 (64-bit)
- Windows Server 2012 and 2012 R2 (64-bit)
- **RAM: 128 MB**
- **Disk space: 124 MB for JRE; 2 MB for Java Update**
- **Processor: Minimum Pentium 2 266 MHz processor**
- Browsers: Internet Explorer 9 and above, Firefox

MAC OS X

- Intel-based Mac running Mac OS X 10.8.3+, 10.9+
- Administrator privileges for installation
- 64-bit browser
- A 64-bit browser (Safari, for example) is required to run Oracle Java on Mac.

LINUX

- Oracle Linux 5.5+
- Oracle Linux 6.x (32-bit), 6.x (64-bit)
- Oracle Linux 7.x (64-bit)2 (8u20 and above)
- Red Hat Enterprise Linux 5.5+1 6.x (32-bit), 6.x (64-bit)
- Red Hat Enterprise Linux 7.x (64-bit)2 (8u20 and above)
- Suse Linux Enterprise Server 10 SP2+, 11.x
- Suse Linux Enterprise Server 12.x (64-bit)2 (8u31 and above)
- Ubuntu Linux 12.04 LTS, 13.x
- Ubuntu Linux 14.x (8u25 and above)
- Ubuntu Linux 15.04 (8u45 and above)
- Ubuntu Linux 15.10 (8u65 and above)
- Browsers: Firefox

PROS & CONS FOR SOLR

Source :

<https://www.realdecoy.com/2018/10/04/solr-powered-search-should-you-look-into-it/>

pros	cons	Weighing
Open source = affordable	The need for external monitoring	need to add monitoring tools - such as SPM and AppDynamics
Large community = active support	Open-source means no guarantees	need to address indexing - ideal memory of 12GB
Scalability = flexibility and growth		open-source
Multi-faceted faceting		

Bottom line

You will need :

- a fully capable development team
- adequate resources in terms of memory and hardware for the release systems
- the right documentation for both the business requirements and the solution itself

SOLR VS. SOLRCLOUD

<https://www.quora.com/What-factors-should-I-take-into-consideration-in-choosing-between-plain-vanilla-Solr-vs-SolrCloud>

- **How large is your data**, does it fit on a single machine and can a single server respond to your queries in a time that is acceptable by your use case? The larger the data, the easier it will be to handle that using SolrCloud, especially when you need to divide it into multiple nodes, handle replication and so on.
- **Is a single Solr node able to index my data** in a acceptable time? If no than the SolrCloud is something that you will probably benefit from as it is easy to distribute a single collection between multiple nodes.
- How near real time you need or want to be? The closer to near real-time the more likely SolrCloud would be a better choice.
- Is Solr a crucial part of the whole application stack - so do we need it to be fault tolerant and highly available? Yes - - SolrCloud is probably a good choice.
- Do I need SolrCloud only features?
- **Is my infrastructure dynamic?** Yes - SolrCloud is probably a good choice because of ease of replica management and autoscaling framework.
- **Will my data grow in foreseeable future and how much it will grow?** If it will grow in a way that it may be a need to distribute - use SolrCloud.

SOLRCLOUD QUESTIONNAIRE

Q1. Standalone mode: Is there a way to identify the Solr source is changed ?

Q1-1. Can Cron job forcibly triggered commit or index rebuild ? (Reindex ?)

Q2. Standalone mode: How to reindex/rebuild 'Core' on Standalone ? (SolrCloud can reindex 'Collection' using API)

<http://.../solr/admin/collections?action=REINDEXCOLLECTION&name=newCollection>

Q3. Master/Slave: If storing the indexes locally on multiple web servers, How can we auto-sync between Solr master and multiple Solr slave other than timer interval setting? Heartbeat or Sync Signal

Q4. Sitecore: Is it possible **RemoteRebuild** to centrally manage for multiple CDs from CM ?

- **onPublishEndAsync strategy** - this event is triggered via the EventQueue to all CDs
 - doing some administrative connection to each CD server (via browser, PowerShell etc) and doing rebuild by just like **RPC Call or RESTFull Service Call**
- (This case we may not need replication setting on Slave Solr or Use both way)

Q5. Solr: How to calculate Solr server required RAM size for Soft commit ?

- By number of page per seconds or Query count per second ?
- How to get this information on Sitecore ?

Q6. Solr: How to check the number of queries per day on Sitecore, Measure the counts per second on peak time ?

(Sitecore initiates query to Solr by user's request)

Q7. Solr: What is Sitecore guideline for the Solr response time per seconds or per query request and Maximum number of connections per server ?

Q8. SolrCloud vs Standalone: On multiple Sharding Solr model, the query response can be faster than single Sharding.

If there is only one Sharding on SolrCloud then what is the benefits compare to Standalone model ? (Only SolrCloud have Reindex Collection)

Q9. Standalone mode: What is minimum Solr standalone server requirements ?

Can Solr be installed on CD or Is there a reason that should not be installed on the same server ? (May be development environment will be acceptable)

Q10. Can Sitecore Docker version be used on Production ?