

ReleaseNote41

January 2013, Apache Solr™ 4.1 available

The Lucene PMC is pleased to announce the release of Apache Solr 4.1

Solr is the popular, blazing fast, open source NoSQL search platform from the Apache Lucene project. Its major features include powerful full-text search, hit highlighting, faceted search, dynamic clustering, database integration, rich document (e.g., Word, PDF) handling, and geospatial search. Solr is highly scalable, providing fault tolerant distributed search and indexing, and powers the search and navigation features of many of the world's largest internet sites.

Solr 4.1 is available for immediate download at:

<http://lucene.apache.org/solr/mirrors-solr-latest-redir.html>

Note: starting with Solr 4.1, the "apache-" prefix has been removed from all artifact and distribution filenames.

See the CHANGES.txt file included with the release for a full list of details.

Solr 4.1 Release Highlights:

SolrCloud enhancements (see <http://wiki.apache.org/solr/SolrCloud>):

- * Simple multi-tenancy through enhanced document routing:
 - The "compositeId" router is the default for collections with hash based routing (i.e. when numShards=N is specified on collection creation).
 - Documents with ids sharing the same domain/prefix, e.g. 'customerB!', will be routed to the same shard, allowing for efficient querying. At query time, one can specify a "shard.keys" parameter that lists the domains, e.g. 'shard.keys=customerB!', and controls what shards the query is routed to.
 - Collections that do not specify numShards at collection creation time use custom sharding and default to the "implicit" router. Document updates received by a shard will be indexed to that shard, unless a "_shard_" parameter or document field names a different shard.
- * Short circuiting for distributed search if a request only needs to query a single shard.
- * Allow creating more than one shard per instance with the Collection API.
- * Allow access to the collections API through CloudSolrServer without referencing an existing collection.
- * Collection API: Support for specifying a list of Solr addresses to spread a new collection across.
- * New and improved auto host detection strategy.
- * Numerous bug fixes and general hardening - it's recommended that all Solr 4.0 SolrCloud users upgrade to 4.1.

New features:

- * The majority of Solr's features, including replication, now work with custom Directory and DirectoryFactory implementations.
- * Indexed term offsets, specifiable via a 'storeOffsetsWithPositions' flag on field definitions in the schema. Useful for highlighters.
- * Solr QParsers may now be directly invoked in the lucene query syntax via localParams and without the _query_ magic field hack.
Example: foo AND {!term f=myfield v=\$qq}
- * Solr now parses request parameters (from URL or sent with POST using content-type application/x-www-form-urlencoded) in its dispatcher code. It no longer relies on special configuration settings in Tomcat or other web containers to enable UTF-8 encoding, which is mandatory for correct Solr behaviour. Solr now works out of the box with e.g. Tomcat, JBoss,...
- * Directory IO rate limiting based on the IO context.
- * Distributed search support for MoreLikeThis.
- * Multi-core: On-demand core loading and LRU-based core unloading after reaching a user-specified maximum number.
- * The new Solr 4 spatial fields now work with the {!geofilt} and {!bbox}

query parsers. The score local-param works too.

- * Extra statistics to RequestHandlers - 5 & 15-minute reqs/sec rolling averages; median, 75th, 95th, 99th, 99.9th percentile request times.
- * PostingsHighlighter support (see <http://blog.mikemccandless.com/2012/12/a-new-lucene-highlighter-is-born.html>)

Admin UI improvements:

- * Internet Explorer is now supported
- * Enhanced readability of XML query response display in Query UI
- * Many improvements to DataImportHandler UI
- * Core creation and deletion now updates the main/left list of cores
- * Admin Cores UI now redirects to newly created core details
- * Deleted documents are calculated/displayed
- * Allow multiple Items to stay open on Plugins-Page

Storage improvements (thanks to the new Lucene 4.1 codec):

- * Faster search, in particular for rare terms such as primary key/id fields.
- * Stored fields are compressed. (See <http://blog.jpountz.net/post/33247161884/efficient-compressed-stored-fields-with-lucene>)

DataImportHandler contrib module backwards-compatibility breaks:

- * These default to the "root" Locale, rather than the JVM default locale as before.
 - NumberFormatTransformer & DateFormatTransformer
 - "formatDate" evaluator
 - "dataimport.properties" file "last_index_time" property
- * These default to UTF-8 encoding, rather than the JVM default encoding as before.
 - FileDataSource & FieldReaderDataSource
- * These may require code changes to custom plug-ins
 - The EvaluatorBag class was eliminated and its public/protected methods were moved to the Evaluator abstract class.
 - The experimental DIHPropertiesWriter interface was renamed DIHProperties, changed to an abstract class and given new signature.

Solr 4.1 also includes numerous optimizations and bugfixes.

Please report any feedback to the mailing lists
(<http://lucene.apache.org/solr/discussion.html>)

Note: The Apache Software Foundation uses an extensive mirroring network for distributing releases. It is possible that the mirror you are using may not have replicated the release yet. If that is the case, please try another mirror. This also goes for Maven access.