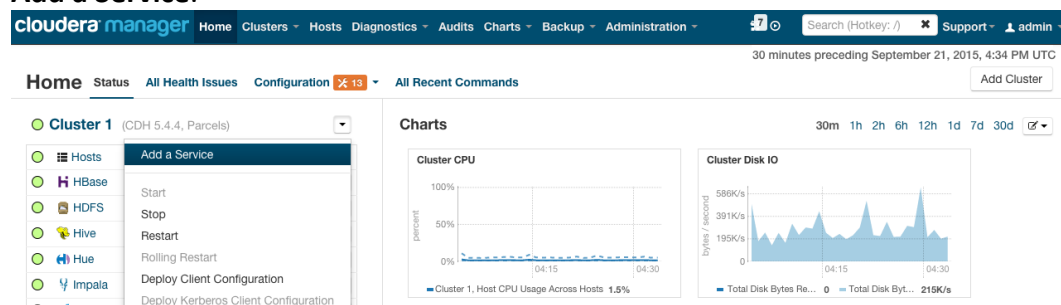




Solr Installation

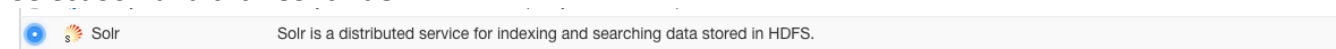
September 2015

HDFS and Zookeeper services are 2 pre-requisites for Solr. Go to Cloudera Manager main page and click **Add a Service.**



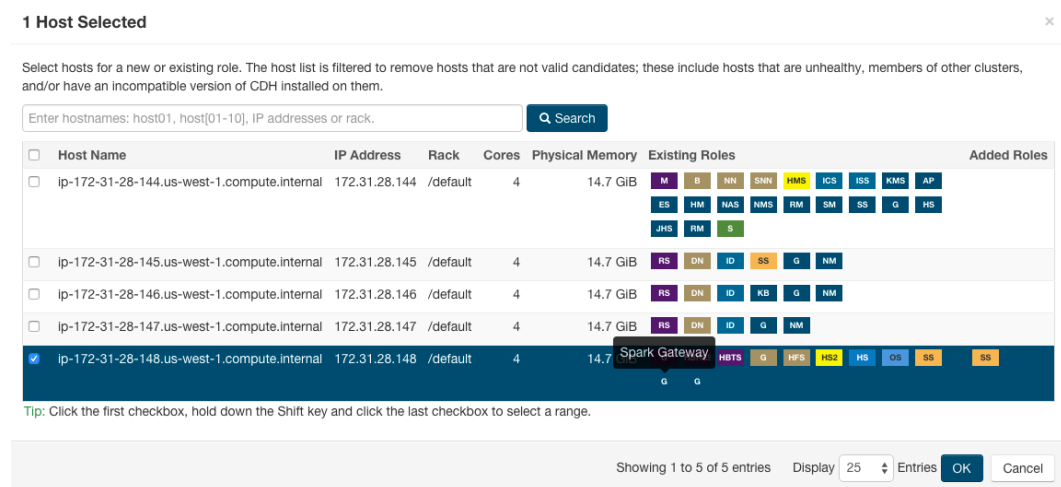
The screenshot shows the Cloudera Manager interface. At the top, there's a navigation bar with 'Home', 'Clusters', 'Hosts', 'Diagnostics', 'Audits', 'Charts', 'Backup', and 'Administration'. Below this, the 'Home' tab is selected, showing 'Status', 'All Health Issues', 'Configuration' (with a red '13' icon), and 'All Recent Commands'. On the left, under 'Cluster 1 (CDH 5.4.4, Parcels)', the 'Add a Service' dropdown menu is open, listing options: Start, Stop, Restart, Rolling Restart, Deploy Client Configuration, and Deploy Kerberos Client Configuration. On the right, there are two charts: 'Cluster CPU' and 'Cluster Disk IO'.

Select Solr and click **Continue.**



The screenshot shows the 'Solr' service selection screen. It features the Solr logo and a description: 'Solr is a distributed service for indexing and searching data stored in HDFS.'

Select HDFS and Zookeeper, and select one or more machines to host Solr server.

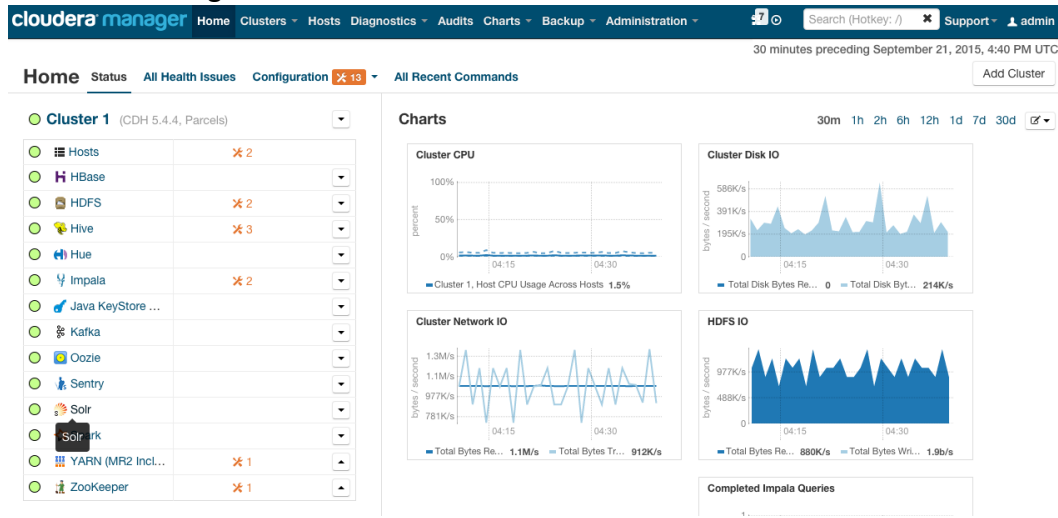


The screenshot shows the '1 Host Selected' dialog. It includes a search bar and a table of hosts. The table has columns for Host Name, IP Address, Rack, Cores, Physical Memory, Existing Roles, and Added Roles. The first host is selected, and its roles are listed as ES, HM, NAS, NMS, RM, SM, SS, G, HS, JHS, RM, S. The second host is also selected, and its roles are listed as RS, DN, ID, SS, G, NM. The third host is selected, and its roles are listed as RS, DN, ID, KB, G, NM. The fourth host is selected, and its roles are listed as RS, DN, ID, G, NM. The fifth host is selected, and its roles are listed as Spark Gateway, HBTS, G, HFS, HSB, HS, OS, SS, SS. A tip at the bottom says: 'Tip: Click the first checkbox, hold down the Shift key and click the last checkbox to select a range.'

Host Name	IP Address	Rack	Cores	Physical Memory	Existing Roles	Added Roles
<input type="checkbox"/> ip-172-31-28-144.us-west-1.compute.internal	172.31.28.144	/default	4	14.7 GiB	M, B, NN, SHN, HNS, JCS, ISS, KMS, AP, ES, HM, NAS, NMS, RM, SM, SS, G, HS, JHS, RM, S	
<input type="checkbox"/> ip-172-31-28-145.us-west-1.compute.internal	172.31.28.145	/default	4	14.7 GiB	RS, DN, ID, SS, G, NM	
<input type="checkbox"/> ip-172-31-28-146.us-west-1.compute.internal	172.31.28.146	/default	4	14.7 GiB	RS, DN, ID, KB, G, NM	
<input type="checkbox"/> ip-172-31-28-147.us-west-1.compute.internal	172.31.28.147	/default	4	14.7 GiB	RS, DN, ID, G, NM	
<input checked="" type="checkbox"/> ip-172-31-28-148.us-west-1.compute.internal	172.31.28.148	/default	4	14.7 GiB	Spark Gateway, HBTS, G, HFS, HSB, HS, OS, SS, SS	

Showing 1 to 5 of 5 entries Display 25 Entries OK Cancel

Cloudera Manager starts Solr servers.



Post verification

Test Solr functionality with Solr utilities and curl tool (Example).

```
./create.sh
```

```
#!/bin/sh
```

```
ZK="172.31.28.144"
```

```
COLLECTION="test"
```

```
echo "create solr collection"
```

```
rm -rf tmp/*
```

```
solrctl -zk $ZK:2181/solrinstancedir --generate tmp/${COLLECTION}_configs
```

```
cp template/schema.xml tmp/${COLLECTION}_configs/conf/
```

```
solrctl -zk $ZK:2181/solrinstancedir --create $COLLECTION tmp/${COLLECTION}_configs
```

```
solrctl -zk $ZK:2181/solr collection --create $COLLECTION -s1-r1
```

```
solrctl -zk $ZK:2181/solr collection --list
```

```
./delete.sh
```

```
#!/bin/sh
```

```
ZK="172.31.28.144"
```

```
COLLECTION="test"
```

```
echo "delete solr collection"
```


```
solrctl -zk $ZK:2181/solr collection --delete $COLLECTION
```

```
solrctl -zk $ZK:2181/solrinstancedir --delete $COLLECTION
```

```
rm -rf tmp/*
```

template/schema.xml
<pre> <?xml version="1.0" encoding="UTF-8" ?> <schema name="example" version="1.5"> <fields> <field name="id" type="string" indexed="true" stored="true" required="true" multiValued="false"/> <field name="doc_name" type="string" indexed="true" stored="true"/> <field name="doc_type" type="string" indexed="true" stored="true"/> <field name="doc_text" type="text_ws" indexed="true" stored="true"/> <field name="_version_" type="long" indexed="true" stored="true"/> </fields> <uniqueKey>id</uniqueKey> <types> <fieldType name="string" class="solr.StrField" sortMissingLast="true"/> <fieldType name="int" class="solr.TrieIntField" precisionStep="0" positionIncrementGap="0"/> <fieldType name="float" class="solr.TrieFloatField" precisionStep="0" positionIncrementGap="0"/> <fieldType name="long" class="solr.TrieLongField" precisionStep="0" positionIncrementGap="0"/> <fieldType name="double" class="solr.TrieDoubleField" precisionStep="0" positionIncrementGap="0"/> <fieldType name="text_ws" class="solr.TextField" positionIncrementGap="100"> <analyzer> <tokenizer class="solr.WhitespaceTokenizerFactory"/> </analyzer> </fieldType> </types> </schema> </pre>
sample/data.json
<pre> [{"id":"doc1", "doc_name":"spark", "doc_type":"word", "doc_text":"alex spark"}, {"id":"doc2", "doc_name":"impala", "doc_type":"pdf", "doc_text":"aleximpala"}] </pre>
./update.sh
<pre> curl -i 'http://172.31.28.148:8983/solr/test/update/json?commit=true' --data-binary @sample/data.json -H 'Content-type:application/json' </pre>
./query.sh
<pre> curl -i 'http://172.31.28.148:8983/solr/test/select?q=%3A*&wt=json&indent=true' </pre>

Enable Kerberos in Solr.

 **Solr** [Status](#) [Instances](#) [Configuration](#) [Commands](#) [Audits](#) [Charts](#) ▼

Configuration Selected Filters: × kerberos × Security × Clear All

Filters

▼ SEARCH

▼ STATUS

All 2

Error 0

Warning 0

Edited 0

Non-default 1

Has Overrides 0

Reason for change...

Save Changes

Kerberos Principal

Solr (Service-Wide)

Solr Secure Authentication

Solr (Service-Wide) ↕

☐ simple

☒ kerberos

Modified version of curl commands to work with Kerberos'edSolr (kinit first before running commands below).

./update.sh
curl -i--negotiate -u : 'http://172.31.28.148:8983/solr/test/update/json?commit=true' --data-binary @sample/data.json -H 'Content-type:application/json'
./query.sh
curl -i --negotiate -u : 'http://172.31.28.148:8983/solr/test/select?q=%3A*&wt=json&indent=true'