Core Discovery (4.4 and beyond)

Discovery-based core enumeration

⚠ old style solr.xml] - there's a new way of defining cores. Core discovery was introduced in 4.3.0 by SOLR-4196, but it is fundamentally broken in all 4.3 versions and not fixed until 4.4. Using the [Solr.xml (supported through 4.x) is recommended unless you have Solr4.4 or higher.

TODO: Lots to flesh out here, this is just barely a start for the docs, consider it a placeholder that will be expanded as time permits. Please ask any questions that come up on the user's list or whatever, and we'll expand this page. Or edit it yourself, all help appreciated!

- Please merge this page here: Solr.xml 4.4 and beyond.

Cores are no longer necessarily defined in in solr.xml. For the near term, solr.xml can exist either in the new or old styles. Whether solr.xml is interpreted as new or old-style is triggered by the absence or presence of a <cores> tag.

- If there is a <cores> element, it is assumed to be an old-style solr.xml.
- If there is *not* a <cores> tag, then this is assumed to be a new-style solr.xml file and the cores are enumerated from SOLR_HOME (i.e. auto discovery).
- In either case, if patterns are detected from both formats, errors are thrown. For example, if you have an old-style solr.xml and you include the str or int tags from the new format, Solr will not start up properly and an error will be logged.

Format for new-style solr.xml

The solr.xml file has a new format, here is a sample. NUMBERS MADE UP FOR NOW!

```
<solr>
  <str name="adminHandler">handler</str>
  <int name="coreLoadThreads">3</str>
  <str name="coreRootDirectory">path to core root</str>
  <str name="managementPath">path</str>
  <str name="sharedLib">lib to share</str>
  <str name="shareSchema">true|false</str> // NOTE, may be replaced by using config sets
  <int name="transientCacheSize">25</int>

  <solrcloud>
    <int name="distribUpdateConnTimeout">30000</int>
    <int name="distribUpdateSoTimeout">15000</int>
    <str name="host">url for host</str>
    <str name="hostContext">solr</str>
    <int name="hostPort">7689</int>
    <int name="leaderVoteWait">15000</int>
    <int name="zkClientTimeout">30000</int>
    <str name="zkHost">url to host</str>
  </solrcloud>

  <logging>
    <str name="class">my.class.for.loggin</str>
    <str name="enabled">true|false</str>
    <watcher>
      <int name="size">12</int>
      <int name="threshold">42</int>
  </logging>

  <shardHandlerFactory class="qualified.path" name="handlerName">
    <int name="connTimeout">15000</int>
    <int name="socketTimeout">15000</int>
  </shardHandlerFactory>
</solr>
```

Finding cores

Exploration of the core tree terminates when a file named core.properties is encountered. Discovery of a file of that name is assumed to define the root of a core. There is no a-priori limit on the depth of the tree. That is, the directories under <SOLR_HOME> are explored until a core.properties file is encountered, and that directory is assumed to be the instanceDir for that core. Nested cores are NOT supported. The core.properties file presently recognizes the following entries:

- name - the name of the core required. If this parameter is not present, the core's name will be the name of the directory containing the core. properties file. The directory structure above that directory is not part of the name.
- config - the configuration file. Defaults to solrconfig.xml

- dataDir - the directory where the index, tlog, etc. are stored. Again, since this is discovery-based, omit this unless you have special needs.
- ulogDir - where the transaction log resides. It may be advantageous to put the transaction lot on a different disk than the index.
- schema - the schema file. Defaults to schema.xml
- shard - the shard ID.
- collection - the collection to which this core belongs
- roles - SolrCloud role definition
- properties - properties file to override core definitions. TBD: This is probably obsolete since we're reading a properties file in the first place. Is there a use case for supporting this now?

- loadOnStartup - [true|false] this core should be loaded and a searcher opened when Solr starts.

- transient - [true|false] this core may be unloaded if the core cache exceeds transientCacheSize (defined in solr.propreties)

- coreNodeName - SolrCloud core node name
- ⚠ 4.6 ⚠ configName - *tentative* a name (expected to be in <SOLR_HOME>/configs that contains all of the configuration files (schema.xml, solrconfig.xml and all supporting files e.g. synonyms.txt). Multiple collections can use the same set of configuration files.