Index Structure (based on Svn REV724)

Created: 23-Apr-2010

Compressed original documents

TF	docid	docid	0	TF	docid	docid	docid	docid
docid	docid	0	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...

Postings lists

terms	CF, DF, offset, length, impact length, local_max, term_pos		
	...		
	term-suffix	term-suffix	...	...
...	..., ..., ..., ..., ..., ..., ...		
	...		
	...	...	...	...

Term block with the same prefix

Term block with the same prefix

Second-level dictionary

terms	term-prefix, pos	term-prefix, pos	term-prefix, pos
	... ...	... ...	... ...

First-level dictionary

footer

Postings list detail

terms_in_nodes · collection_frequency · document_frequency · docid_pos_on_disk · impacted_length · postings_len · local_max_impact · string_position

4	4	4	8	4	4	1	4

In case if SPECIAL_COMPRESSION is used,
(1) `docid_pos_on_disk` stores the first postings(4 bytes) in the higher order of the 8 bytes and store the first term frequency (4 bytes) in the lower order of the 8 bytes
(2) `impacted_legnth` stores the second postings (if there is one)
(3) `postings_len` stores the second term frequency (if there is one)

First-level dictionary detail

terms_in_root · term_string · disk_pos_in_second_level

4	4	8

Footer detail

disk_pos_first_level_dict · longest_postings_size · length_of_the_leongest_term · highest_df

8	4	4	8