Index Structure (based on Svn REV1387, ATIRE V0.4)

Created: 26-Jan-2012

impact_value_start

doc_count_start

impact_offset_start

uncompressed impact header

uncompressed postings

| 8 | 8 | 4 | 4 | compressed header | Compressed postings |

postings_chain

chain_length

the_quantum_count

beginning_of_the_postings

terms_in_nodes

document_frequency

local_max_impact

| 4 | 4 | 4 | 8 | 4 | 4 | 1 | 4 |

collection_frequency

docid_pos_on_disk

impacted_length

postings_len

string_position

Compressed original documents

...	...	...	...	...	...	...	...	...	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

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_length stores the second postings (if there is one)
(3) postings_len stores the second term frequency (if there is one)

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

First-level dictionary

footer

terms_in_root

| 4 | 4 | 8 |

term_string

disk_pos_in_second_level

disk_pos_first_level_dict

longest_postings_size

| 8 | 4 | 4 | 8 | 8 | 8 |

length_of_the_leongest_term

highest_df

checksum

collection_id