

Inverted Index Example

Suppose we have a posting list with 4 entries:

[{ docId = 1, freq = 3}, { docId = 258, freq = 4}, { docId = 4,294,967,038, freq = 16,777,215 }]

Group Varint encoding format

(between 5 and 17 bytes used for every 4 digits/every 2 posting list elements)

Simplified view of dictionary:

Posting Pointers	Freq	More columns...
...
	4	...

Length Byte of the form: $L_1 L_2 L_3 L_4$				
L_1 bytes	L_2 bytes	L_3 bytes	L_4 bytes	
Number 1	Number 2	Number 3	Number 4	

Where L_i is 2 bits, representing length of integer(between 1 and 4 bytes)

Postings file:

...
Num (1B) 00010000	Num (1B) 1	Num (1B) 3	Num (2B) 257	Num (1B) 4	Num (1B) 11100000
Num (4B) 4,294,966,780	Num (3B) 16,777,215	Num (1B) 0	Num (1B) 0