

Flat Bloofi

Soluzioni "flat" per Bloom Filters multidimensionali



UNIVERSITÀ DEGLI STUDI
DI SALERNO

Dipartimento di Informatica

Relatore

Ch.ma prof.ssa
Rosalba Zizza

Candidato

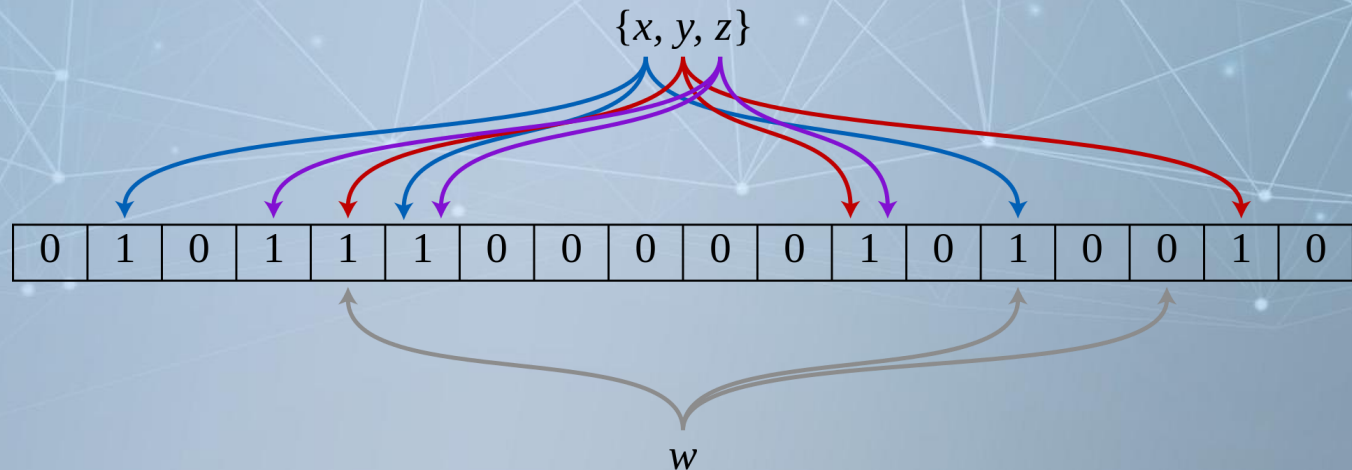
Pierluigi Liguori
Mat. 0512104224

a.a. 2019/2020



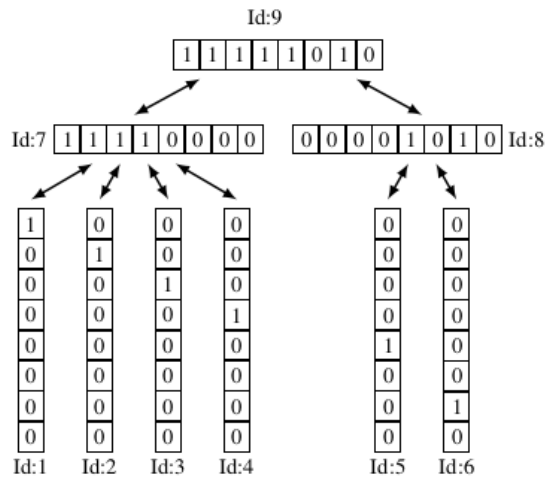
Bloom Filter

- m parole da inserire
- Array di N elementi
- K funzioni hash
- Probabilità di falsi positivi

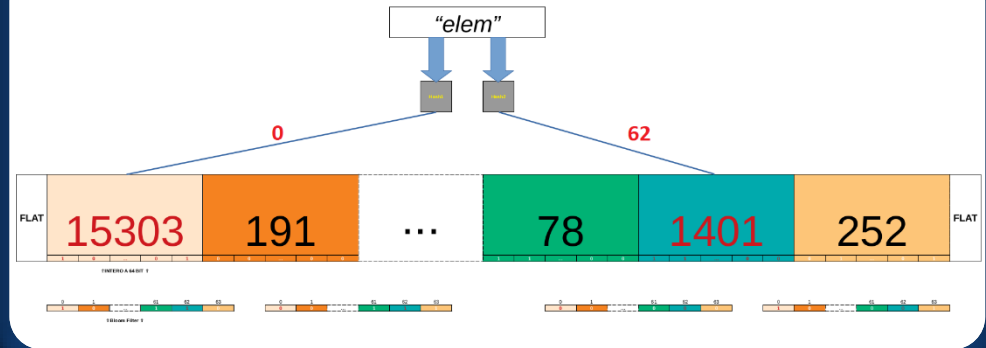


BLOOFI

- Bloofi

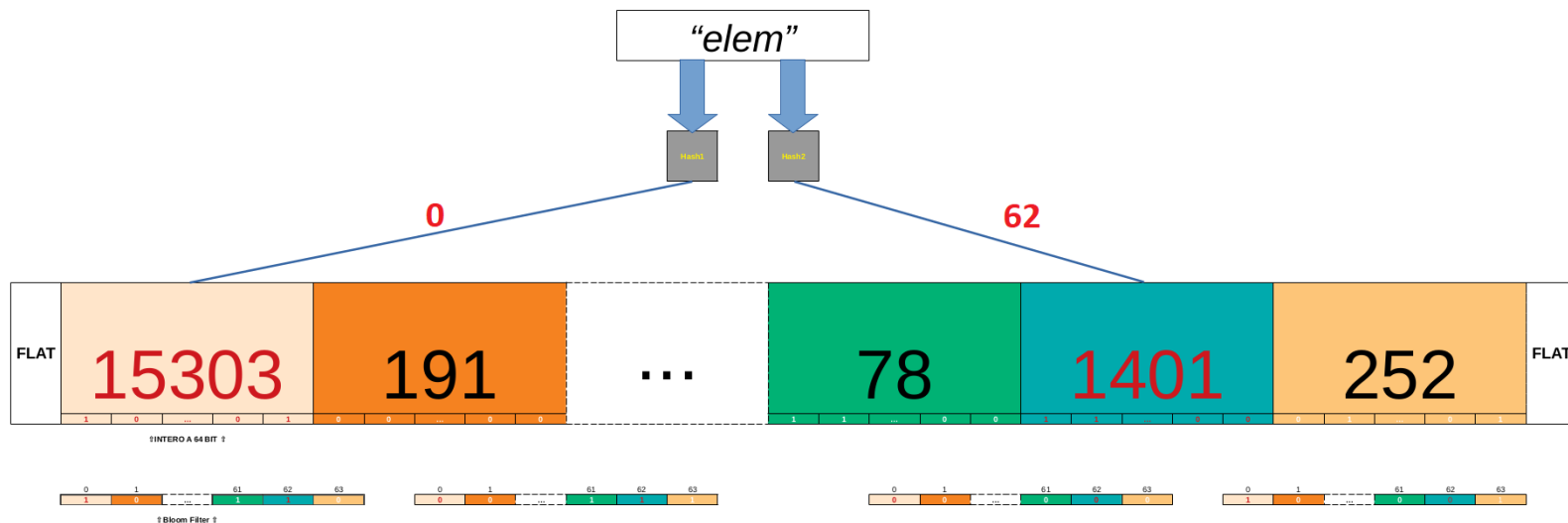


- Flat Bloofi

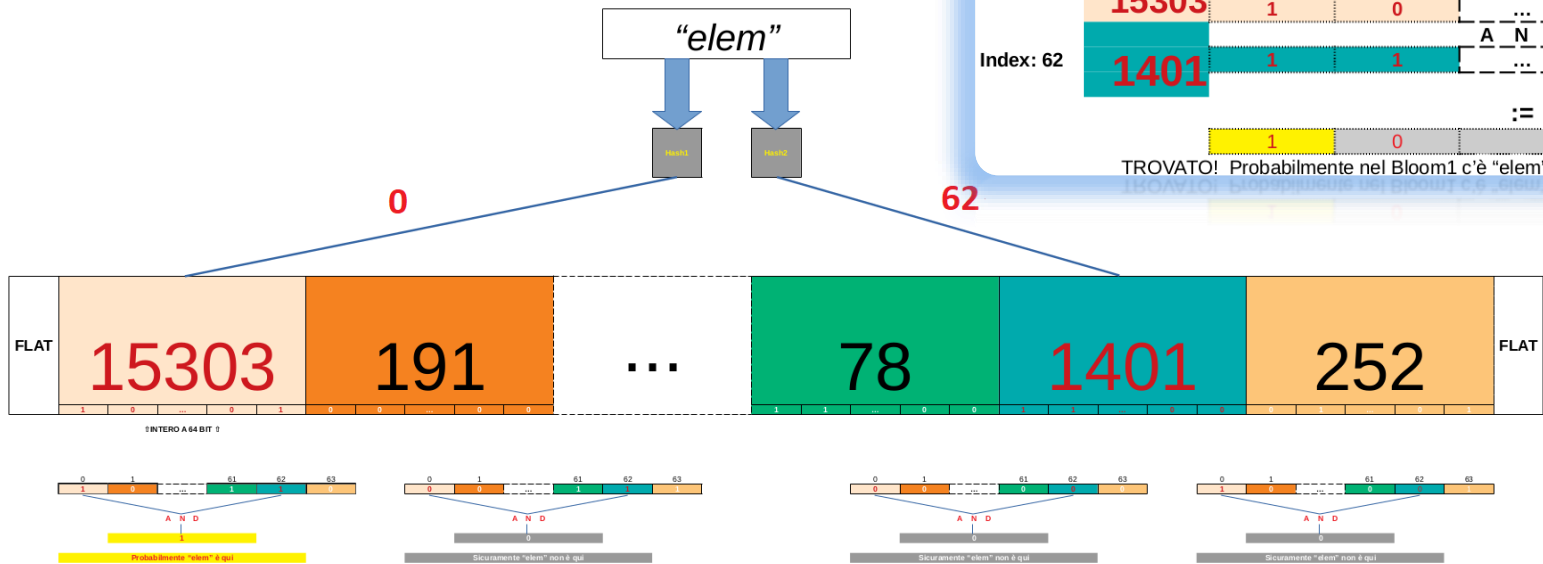


Flat Bloofi

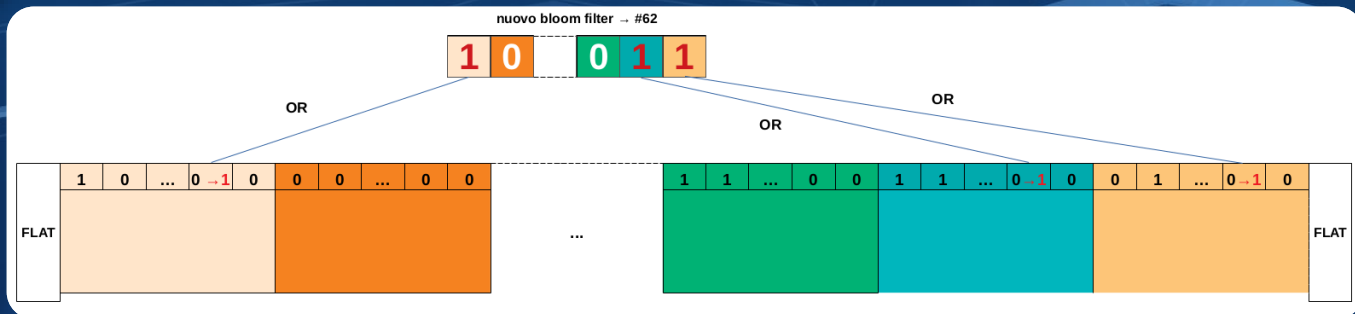
- Z Flat Bloofi ($Z \times 64$ Bloom Filters)
- array β di dimensione $Z \times 64$
- tabella hash



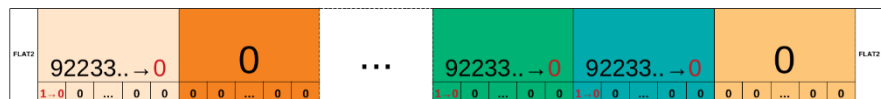
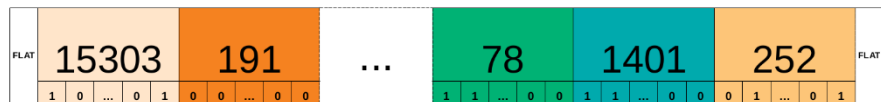
Operazione di ricerca



Inserimento

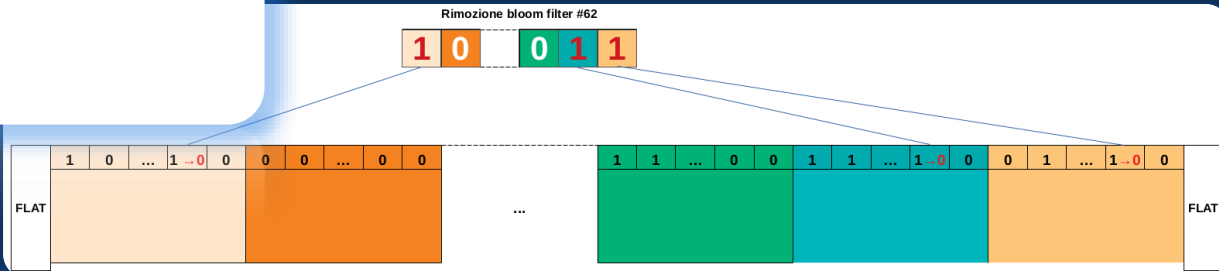


Rimozione (caso b.)



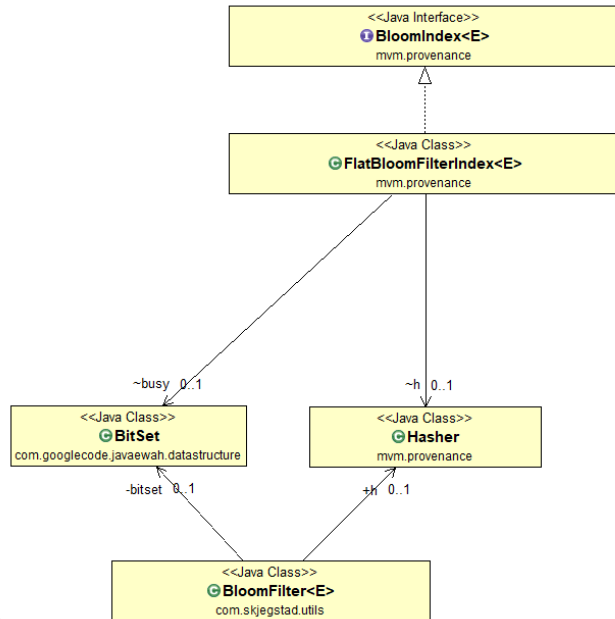
Rimozione Bloom filter #64

Rimozione (caso a.)



Implementazioni

- Java



Anomalia remove in Java

```
7 public static void main(String[] args) {
8     Hasher h = new Hasher();
9     BloomIndex<Integer> flatbfi = new FlatBloomFilterIndex<Integer>();
10
11     //Il primo flat viene riempito con 64 bloom filter.
12     for (int i = 0; i < 64; i++) {
13         BloomFilter<Integer> bf = new BloomFilter<Integer>(h, 0.01, 100, 1);
14         flatbfi.insertBloomFilter(bf, null);
15     }
16     System.out.println("Bloom filter aggiunti: " + flatbfi.getSize());
17     //Aggiungo un nuovo bloom filter
18     BloomFilter<Integer> bf01 = new BloomFilter<Integer>(h, 0.01, 100, 1);
19     flatbfi.insertBloomFilter(bf01, null);
20     System.out.println("Bloom filter aggiunti: " + flatbfi.getSize());
21
22     flatbfi.deleteFromIndex(65, null);
23 }
24
25 @Override
26 public int deleteFromIndex(int id, InsDelUpdateStatistics stat) {
27     int index = idMap.remove(id);
28     idMap.remove(id);
29     busy.unset(index);
30     if (busy.getWord(index / 64) == 0) {
31         System.out.println("mi tocca eliminare un intero flat ...");
32         for (int k = index / 64 * 64; k < index / 64 * 64 + 64; ++k)
33             fromIndexId.remove(k);
34         buffer.remove(index / 64);
35         busy.removeWord(index / 64);
36         for (Map.Entry<Integer, Integer> me : idMap.entrySet()) {
37             if (me.getValue().intValue() / 64 == index / 64) {
38                 idMap.put(me.getKey(), me.getValue().intValue() - 64);
39             }
40         }
41     } else {
42         clearBloomAt(index);
43     }
44 }
```

Problems Javadoc Declaration Console

terminated: provaBugRemove [Java Application] /usr/lib/jvm/java-8-openjdk-amd64/bin/java (16 mag 2020, 00:11:12)

#bloom filter aggiunti:64
#bloom filter aggiunti:65
mi tocca eliminare un intero flat ...
Exception in thread "main" java.lang.IndexOutOfBoundsException: Index: 65, Size: 64
at java.util.ArrayList.rangeCheck(ArrayList.java:657)
at java.util.ArrayList.remove(ArrayList.java:496)
at mvm.provenance.FlatBloomFilterIndex.deleteFromIndex(FlatBloomFilterIndex.java:33)

- C

```
struct flatbloofi{

    list fromindextoId;
    struct DataItem* idMap[SIZE];
    list buffer;
    bitset_t* busy;

    struct hasher* h;

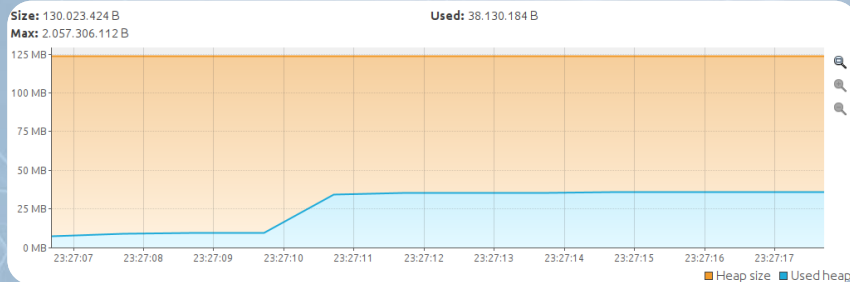
};
```


Testing

- Test di correttezza
- Analisi memoria
- Tempi di esecuzione

Java – 38 Mb

C – 25 Mb



```

bfList=insertElement(bfList,bf); //Inserimento del bf nella lista
bfList=insertElement(bfList,bf2); //Inserimento del bf nella lista
bfList=insertElement(bfList,bf3); //Inserimento del bf nella lista
bfList=insertElement(bfList,bf4); //Inserimento del bf nella lista

for(int i = 0; i < 500; i++) {
    bloom_add(bf,concatena2("gatto", i));
    bloom_add(bf,concatena2("cane", i));

    bloom_add(bf2, concatena2("giacomo", i));
    bloom_add(bf2, concatena2("aldo", i));
    bloom_add(bf2, concatena2("giovanni", i));

    bloom_add(bf3, concatena2("pentola", i));
    bloom_add(bf3, concatena2("forchetta", i));

    bloom_add(bf4, concatena2("pc", i));
}

struct bloom* current = NULL;

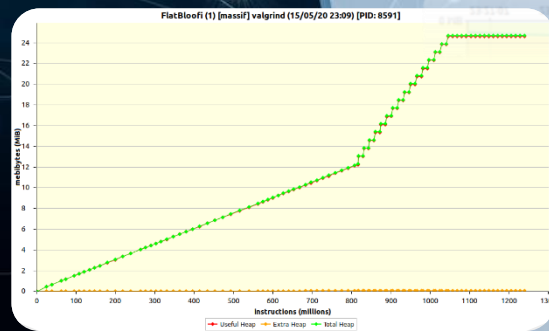
for(int i = 0; i < bfList->size; i++){
    current = getElement(bfList,i);
    insertBloomFilter(bl,current);
}

search(bl, "forchettal77");
    
```

Console:

```

rminated> FlatBloomi [C/C++ Application] /home/fernet/workspace/C/FlatBloomi/Debug/Flat
***** FLAT BLOOMI C *****
parola si trova nel bloom filter con id: 3
    
```

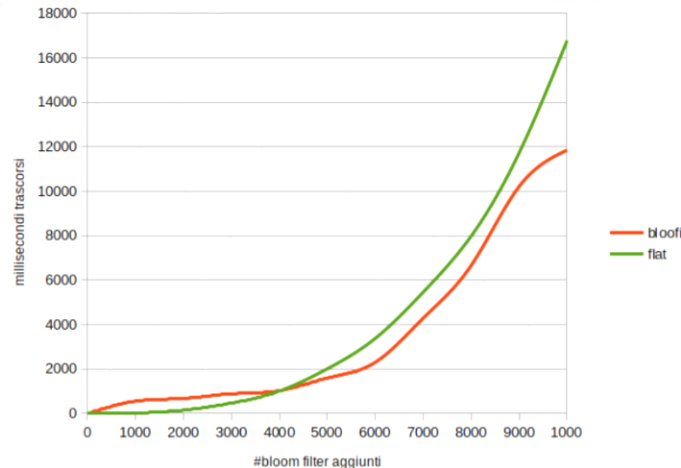


	1 Bloom Filters 5 Elementi ciascuno			100 Bloom Filters 50 Elementi ciascuno			1000 Bloom Filters 500 Elementi ciascuno		
	Inserimento	Aggiornamento	Ricerca	Inserimento	Aggiornamento	Ricerca	Inserimento	Aggiornamento	Ricerca
Java	0,461 ms	0,041 ms	0,025 ms	4,327 ms	0,853 ms	1,597 ms	45,145 ms	14,734 ms	27,320 ms
C	0,120 ms	0,008 ms	0,005 ms	3,220 ms	0,986 ms	0,309 ms	67,871 ms	30,1992 ms	22,089 ms

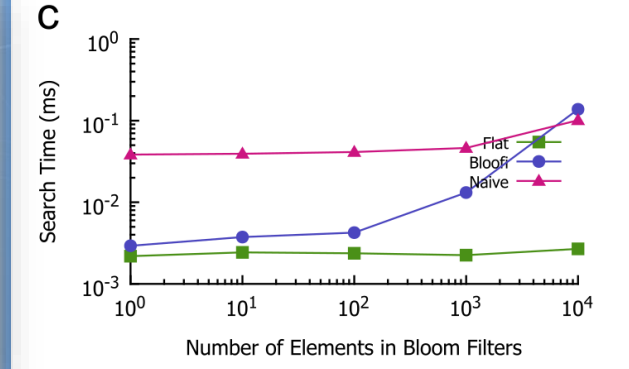
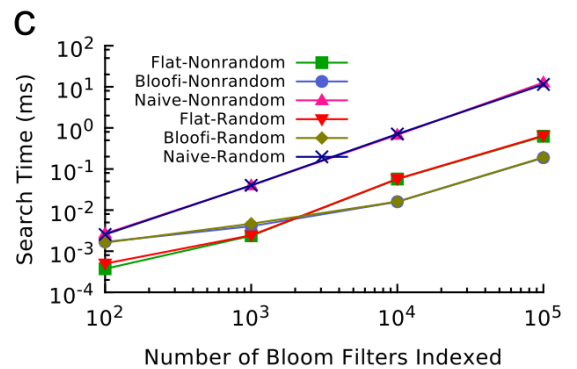
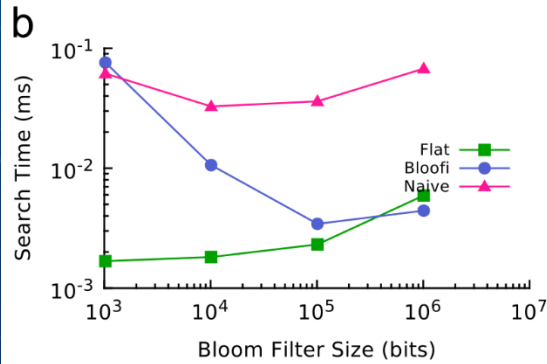
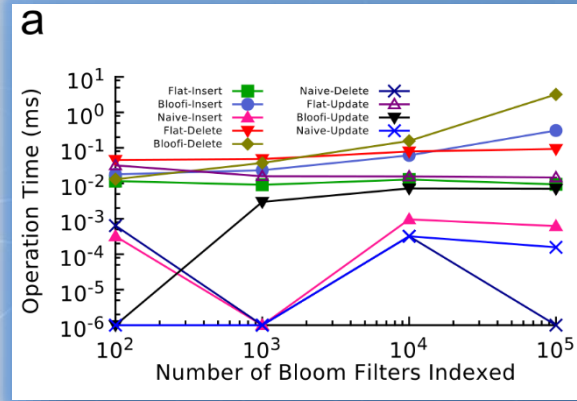
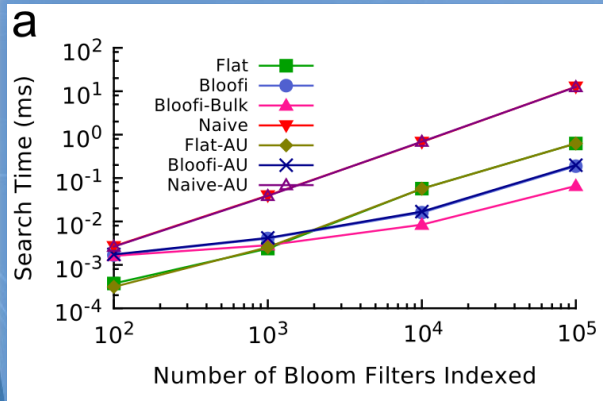
*nextSetBit si può migliorare

CFlatBloofi VS CBloofi

	1 Bloom Filters 5 Elementi ciascuno			100 Bloom Filters 50 Elementi ciascuno			1000 Bloom Filters 500 Elementi ciascuno		
	Inserimento	Aggiornamento	Ricerca	Inserimento	Aggiornamento	Ricerca	Inserimento	Aggiornamento	Ricerca
Cbloofi	0.011ms	0.011 ms	0.002ms	13.385 ms	2.453 ms	4.087 ms	162.464 ms	30.663 ms	549.82 ms
Cflat	0,120 ms	0,008 ms	0,005 ms	3,220 ms	0,986 ms	0,309 ms	67,871 ms	30,1992 ms	22,089 ms



Variazione dei parametri





Grazie per l'attenzione