Analysis of Comparison of SIMC Indexing Methods

Chang GUO, Libo ZHUO

For a range of parameter settings:

- size of tuples (determined by number of attributes)
- size of bit-slices (determined by expected number of tuples)
- size of tuple and page signatures (determined by pF)

In this project, we check and analysis comparison of parameters via following queries
- not using indexing (use x as the signature type)
- open query (?,?,...,?)
- query with one solution (100001,?,...,?)
- query with many solutions (?,?,a3-001,...,?)
- query with multiple values (100001,?,a3-001,...,?)

Each query with its implementation and shell responses as follows, it is obviously that bit slice is the fastest way to index.
---------------- Test Running... ----------------

---------------- Building Database... ----------------

---------------- Generating Data... ----------------
1000000,NwLRbBMqbHcdARZoWkKy,a3-000,a4-000
1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
1000002,bLDBefsaRcbYNeCDYgGX,a3-002,a4-002
1000003,XPKlOrELLnmpAPQFwKho,a3-003,a4-003
1000004,pKmcOqhNwNKueWhSQMgB,a3-004,a4-004
---------------- Creating Relationship... ----------------
---------------- Checking Status... ----------------
Global Info:
Dynamic:
　　#items:　tuples: 0　tsigs: 0　psigs: 0　bsigs: 0
　　#pages:　tuples: 1　tsigs: 1　psigs: 1　bsigs: 2
Static:
　　tups　　#attrs: 4　size: 42 bytes　max/page: 97
　　sigs　　bits/attr: 9
　　tsigs　size: 64 bits (8 bytes)　max/page: 511
　　psigs　size: 5584 bits (698 bytes)　max/page: 5
　　bsigs　size: 8 bits (1 bytes)　max/page: 4092
---------------- Inserting Data... ----------------
Inserting: 1000000,NwLRbBMqbHcdARZoWkKy,a3-000,a4-000
Inserting: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001

Inserting: 1000002,bLDBefsaRcbYNeCDYgGX,a3-002,a4-002
Inserting: 1000003,XPKlOrELLnmpAPQFwKho,a3-003,a4-003
Inserting: 1000004,pKmcOqhNwNKueWhSQMgB,a3-004,a4-004
---------------- Checking Status... ----------------
Global Info:
Dynamic:
  #items: tuples: 5   tsigs: 5   psigs: 1   bsigs: 5584
  #pages: tuples: 1   tsigs: 1   psigs: 1   bsigs: 2
Static:
  tups   #attrs: 4   size: 42 bytes   max/page: 97
  sigs   bits/attr: 9
  tsigs   size: 64 bits (8 bytes)   max/page: 511
  psigs   size: 5584 bits (698 bytes)   max/page: 5
  bsigs   size: 8 bits (1 bytes)   max/page: 4092
---------------- Showing All Data... ----------------
1000000,NwLRbBMqbHcdARZoWkKy,a3-000,a4-000
1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
1000002,bLDBefsaRcbYNeCDYgGX,a3-002,a4-002
1000003,XPKlOrELLnmpAPQFwKho,a3-003,a4-003
1000004,pKmcOqhNwNKueWhSQMgB,a3-004,a4-004
---------------- Executing Query... ----------------
// not using indexing (use x as the signature type)
// here we can see that it examined 5 tuples within one data page
// with no output and matched tuple
./select   R   1000001,?,a3-002,?   x
Query Stats:
signatures read: 0
sig pages read: 0
tuples examined: 5
data pages read: 1
false match pages: 1

// this part we test open query
// expected output is to show all the data in the database
---------------- open query ---- Executing Query... ----------------5

./select R ?,?,?,? t
Matched Tuple: 1000000,NwLRbBMqbHcdARZoWkKy,a3-000,a4-000
Matched Tuple: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
Matched Tuple: 1000002,bLDBefsaRcbYNeCDYgGX,a3-002,a4-002
Matched Tuple: 1000003,XPKlOrELLnmpAPQFwKho,a3-003,a4-003
Matched Tuple: 1000004,pKmcOqhNwNKueWhSQMgB,a3-004,a4-004
Query Stats:
signatures read: 5

sig pages read: 1
tuples examined: 5
data pages read: 1
false match pages: 0
--------------- Executing Query... ----------------
./select R ?,?,?,? p
Matched Tuple: 1000000,NwLRbBMqbHcdARZoWkKy,a3-000,a4-000
Matched Tuple: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
Matched Tuple: 1000002,bLDBefsaRcbYNeCDYgGX,a3-002,a4-002
Matched Tuple: 1000003,XPKIOrELLnmpAPQFwKho,a3-003,a4-003
Matched Tuple: 1000004,pKmcOqhNwNKueWhSQMgB,a3-004,a4-004
Query Stats:
signatures read: 1
sig pages read: 1
tuples examined: 5
data pages read: 1
false match pages: 0
--------------- Executing Query... ----------------
./select R ?,?,?,? b
Matched Tuple: 1000000,NwLRbBMqbHcdARZoWkKy,a3-000,a4-000
Matched Tuple: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
Matched Tuple: 1000002,bLDBefsaRcbYNeCDYgGX,a3-002,a4-002
Matched Tuple: 1000003,XPKIOrELLnmpAPQFwKho,a3-003,a4-003
Matched Tuple: 1000004,pKmcOqhNwNKueWhSQMgB,a3-004,a4-004
Query Stats:
signatures read: 0
sig pages read: 0
tuples examined: 5
data pages read: 1
false match pages: 0
--------------- one solution ---- Executing Query... ----------------
./select R 1000001,?,?,? t
Matched Tuple: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
Query Stats:
signatures read: 5
sig pages read: 1
tuples examined: 5
data pages read: 1
false match pages: 0
--------------- Executing Query... ----------------
./select R 1000001,?,?,? p
Matched Tuple: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
Query Stats:
signatures read: 1

```
# sig pages read:       1
# tuples examined:      5
# data pages read:      1
# false match pages: 0
--------------- Executing Query... ----------------
./select R 1000001,?,?,? b
Matched Tuple: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
Query Stats:
# signatures read:      0
# sig pages read:       0
# tuples examined:      5
# data pages read:      1
# false match pages: 0
--------------- many solutions ---- Executing Query... ---------------
./select R ?,?,?,? t
Matched Tuple: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
Query Stats:
# signatures read:      5
# sig pages read:       1
# tuples examined:      5
# data pages read:      1
# false match pages: 0
--------------- Executing Query... ----------------
./select R ?,?,?,? p
Matched Tuple: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
Query Stats:
# signatures read:      1
# sig pages read:       1
# tuples examined:      5
# data pages read:      1
# false match pages: 0
--------------- Executing Query... ----------------
./select R ?,?,?,? b
Matched Tuple: 1000001,hidDQscdxRjMowfRXSJY,a3-001,a4-001
Query Stats:
# signatures read:      0
# sig pages read:       0
# tuples examined:      5
# data pages read:      1
# false match pages: 0
--------------- multiple values ---- Executing Query... ---------------
./select R 1000001,?,a3-002,? t
Query Stats:
# signatures read:      5
```

```
# sig pages read:      1
# tuples examined:     0
# data pages read:     0
# false match pages: 0
---------------- Executing Query... ----------------
./select R 1000001,?,a3-002,? p
Query Stats:
# signatures read:     1
# sig pages read:      1
# tuples examined:     5
# data pages read:     1
# false match pages: 1
---------------- Executing Query... ----------------
./select R 1000001,?,a3-002,? b
Query Stats:
# signatures read:     0
# sig pages read:      0
# tuples examined:     5
# data pages read:     1
# false match pages: 1
```