dbTouch
ANALYTICS AT YOUR FINGERTIPS

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)

(CIDR 2013 & ICDE 2014)
db kernels for touch-based interactive exploration

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
(big) data is everywhere

CHALLENGE: EASY DATA MANAGEMENT
not always sure what we are looking for (until we find it)
db systems are too heavy and blind!

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
adaptive indexing:
- cracking (CIDR 07)
- cracking updates (SIGMOD 07)
- sideways cracking (SIGMOD 09)
- adaptive indexing (PVLDB 11)
- concurrency control (PVLDB 12)
- stochastic cracking (PVLDB 12)
- adaptive time series indexing (SIGMOD 14)

adaptive loading/code
- basic vision (CIDR 11)
- nodb (SIGMOD 12)
- h20 (SIGMOD 14)

TREAT QUERIES AS ADVICE ON HOW TO STORE AND ACCESS DATA

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
adaptive indexing:
cracking (CIDR 07)
cracking updates (SIGMOD 07)
sideways cracking (SIGMOD 09)
adaptive indexing (PVLDB 11)
concurrency control (PVLDB 12)
stochastic cracking (PVLDB 12)
adaptive time series indexing (SIGMOD 14)

adaptive loading/code
basic vision (CIDR 11)
nodb (SIGMOD 12)
h20 (SIGMOD 14)

TREAT QUERIES AS ADVICE ON HOW TO STORE AND ACCESS DATA

Queries?
just touch the data

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
just touch the data

this is not about query building
it is about query processing
**DEMO** (ελπίζω να δουλέψει - συνήθως δουλεύει)

an *exploration* tool

get a quick feeling about your data

focus on interesting areas
dbTouch

column 1

start

column 2

column 3

query

data

02
72
22
37
70
57
43
what does this mean for db kernels?
select R.a from R
limit
sample
db

select R.a from R
limit
sample

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
select R.a from R
limit
sample
db

select R.a from R
limit
tsample

$\text{dbTouch}$

56 38 45 2

process only what you touch

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
from touch to query processing
from touch to query processing

touch location (x)
from touch to query processing

touch location \((x)\)

size(width)
from touch to query processing

storage layer
from touch to query processing

touch location \((x)\)

size(width)

storage layer

tuples

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
from touch to query processing

storage layer

row ID = \(\frac{\text{tuples} \times x}{\text{size}}\)

touch location \(x\)

size(width)

tuples

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
from touch to query processing

touch location \((x)\)

storage layer

row ID = \((\text{tuples} \times x) / \text{size}\)

tuples
explore: touch, observe and react
explore: touch, observe and react

the system does not have control of the data flow anymore

the user dictates which is the next tuple
get_next()
enable interactive processing
by instant reactions

caching    prefetching    pre-computation

exploit idle time    guess next touch (pos+speed)
enable interactive processing by instant reactions

caching  prefetching  pre-computation

exploit idle time  guess next touch (pos+speed)
enable interactive processing by instant reactions

caching   prefetching   pre-computation

exploit idle time  guess next touch (pos+speed)

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
select avg(R.c) where R.a = S.b and S.b < 20

avg(R.c)

R.a = S.b

σ(<20)

S.b          R.a
select \( \text{avg}(R.c) \) where \( R.a = S.b \) and \( S.b < 20 \)

\[
\begin{align*}
\text{S.b} \\
\text{R.c} \\
\text{R.a}
\end{align*}
\]
select avg(R.c)
where R.a=S.b
and S.b<20
select avg(R.c) where R.a=S.b and S.b<20
select avg(R.c)
where R.a=S.b
and S.b<20
select \( \text{avg}(R.c) \) where \( R.a = S.b \) and \( S.b < 20 \)

**exploration**  
no need for exact/correct answers  
thin/incomplete hash tables
exploreation

no need for exact/correct answers
thin/incomplete hash tables

multiple passes
exploration

no need to access all data all the time
exploration

no need to access all data all the time

visual objects
exploration

no need to access all data all the time

visual objects
exploration

no need to access all data all the time

visual objects

samples hierarchy
exploration

no need to access all data all the time
exploration

no need to access all data all the time

visual objects

samples hierarchy

base data

initial sample
exploration

no need to access all data all the time

visual objects

samples hierarchy

base data

initial sample

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
exploration

no need to access all data all the time

visual objects

samples hierarchy

create incrementally and on demand

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
touch the schema

table
touch the schema
touch the schema

table

from rows to columns
touch the schema

table

from rows to columns

incremental adaptation

Stratos Idreos (Harvard) and Erietta Liarou (EPFL)
design db kernels for touch-based exploration
continuous human-computer interaction
design db kernels for touch-based exploration
continuous human-computer interaction
design db kernels for touch-based exploration
continuous human-computer interaction

Thank you!