Logistics
Discussion papers to be released next week.
Registration for presentation: mid-next week.
Logistics
Discussion papers to be released next week.
Registration for presentation: mid-next week.

**Systems Project should start now:**
Read, design, go to labs, basic tools.
Dedicated section next week - one in person and one online.
Logistics
Discussion papers to be released next week. Registration for presentation: mid-next week.

**Systems Project should start now:**
Read, design, go to labs, basic tools. Dedicated section next week - one in person and one online.

**Research lecture schedule:**
Feb 14: Learned Systems (Subarna)
Feb 16: Neural Networks (Sanket)
Feb 21: Image Storage (Utku)
Feb 23: Blockchain (Hao)
Registration form for preference for project is out.
NoSQL Key-value Stores

machine learning, social media, smart homes, web browsers, phones, web-based apps, security, health devices, graphs, analytics
bloom filters
[min-max] /page
fence pointers

[0,0,1,1,1]
hash fun.

buffer

Level 1

Level 2

Level 3

... Level N

MEMORY

DISK

pages

SSTables

tiered

leveled

sorted

[0,1,0,0,1,1,1]
hash fun.

hash fun.
fundamental building blocks

properties when combined
DESIGN SPACE

COST SYNTHESIS

WHAT-IF
what should you be doing?

READING
Readings for this week (and systems project)


How to read:

review and slides should focus on

what is the problem
why is it important
why is it hard
why existing solutions do not work
what is the core intuition for the solution
solution step by step
does the paper prove its claims
exact setup of analysis/experiments
are there any gaps in the logic/proof
possible next steps

* follow a few citations to gain more background
How to read:

- review and slides should focus on
  - what is the problem
  - why is it important
  - why is it hard
  - why existing solutions do not work
  - what is the core intuition for the solution
  - solution step by step
  - does the paper prove its claims
  - exact setup of analysis/experiments
  - are there any gaps in the logic/proof
  - possible next steps

* follow a few citations to gain more background

How long should it take?
1 week would be great
Today

Advanced LSM-tree/Storage designs. We will see that “advanced is simple”
Today

Advanced LSM-tree/Storage designs. We will see that “advanced is simple”

System project: quick first intro
Today

Advanced LSM-tree/Storage designs.
We will see that “advanced is simple”

System project: quick first intro

How to search the design space: Step 1: cost synthesis.
BITS PER ENTRY IN FILTERS: OPTIMIZED OUT

Monkey: Optimal Navigable Key-Value Store

@SIGMOD2017
BITS PER ENTRY IN FILTERS: OPTIMIZED OUT

Monkey: Optimal Navigable Key-Value Store

@SIGMOD2017
BITS PER ENTRY IN FILTERS: OPTIMIZED OUT

Monkey: Optimal Navigable Key-Value Store

@SIGMOD2017
BITS PER ENTRY IN FILTERS: OPTIMIZED OUT

Monkey: Optimal Navigable Key-Value Store

worst lookup cost:
sum of false positive rates

bits per entry:
fixed per run

buffer

Level 1

Level 2

...

Level N
BITs Per Entry in Filters: Optimized Out

Level 1

Level 2

…

Level N

@SIGMOD2017

the same memory budget is more impactful at smaller levels
BITS PER ENTRY IN FILTERS: OPTIMIZED OUT

Monkey: Optimal Navigable Key-Value Store

the same memory budget is more impactful at smaller levels

buffer

Level 1

Level 2

Level N
BITS PER ENTRY IN FILTERS: OPTIMIZED OUT

Monkey: Optimal Navigable Key-Value Store

@SIGMOD2017

the same memory budget is more impactful at smaller levels

lookup cost

update cost

WiredTiger

Cassandra, HBase

RocksDB, LevelDB

monkey

buffer

Level 1

Level 2

Level N
The same memory budget is more impactful at smaller levels.
MERGE POLICY: SHOULD BE TUNED

Dostoevsky: Space-Time Optimized Evolvable Scalable Key-Value Store
MERGE POLICY: SHOULD BE TUNED

Dostoevsky: Space-Time Optimized Evolvable Scalable Key-Value Store
merge policy: fixed across levels

MERGE POLICY: SHOULD BE TUNED

Dostoevsky: Space-Time Optimized Evolvable Scalable Key-Value Store
MERGE POLICY: SHOULD BE TUNED

Dostoevsky: Space-Time Optimized Evolvable Scalable Key-Value Store

merge policy: fixed across levels

- Level 1
- Level 2
- ...
- Level N
merging small levels does not help that much (point, range, space)

merge policy: fixed across levels

buffer

Level 1

Level 2

... 

Level N
merging small levels does not help that much (point, range, space)
merging small levels does not help that much (point, range, space)
merging small levels does not help that much (point, range, space)

buffer

Level 1

Level 2

... Level N

MERGE POLICY: SHOULD BE TUNED

Dostoevsky: Space-Time Optimized Evolvable Scalable Key-Value Store

@SIGMOD2018

Dostoevsky

Monkey

RocksDB well tuned

RocksDB

normalized throughput (ops/s)

1.0

0.9

0.8

0.7

0.6

0.5

0.4

0.3

0.2

0.1

0.0

10^{-2}

10^{-1}

read/write
Stratos Idreos, Konstantinos Zoumpatianos, Brian Hentschel, Michael Kester, Demi Guo. In Proceedings of the ACM SIGMOD International Conference on Management of Data, 2018

+ Technical Report if you want to see design primitives in more detail
BIG DATA SYSTEMS

NoSQL | Neural Networks | SQL | Graph | Data Science