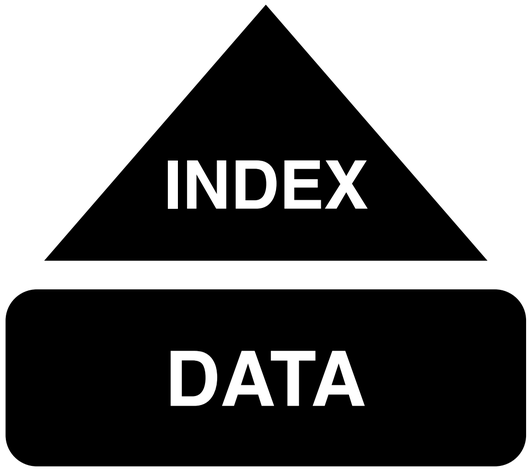
CS 265

Stratos Idreos

BIG DATA SYSTEMS

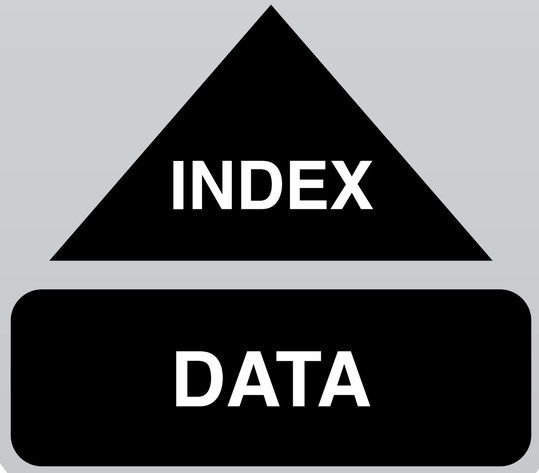
NoSQL | Neural Networks | SQL | Graph | Data Science



HOW
TO STORE
DATA

ALGORITHMS

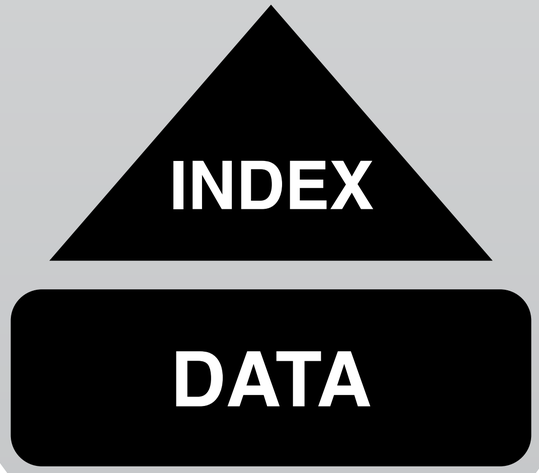
data structure decisions define
the algorithms that access data



ALGORITHMS

unordered

[7,4,2,6,1,3,9,10,5,8]



ALGORITHMS

unordered

[7,4,2,6,1,3,9,10,5,8]

INDEX

DATA

ALGORITHMS

unordered

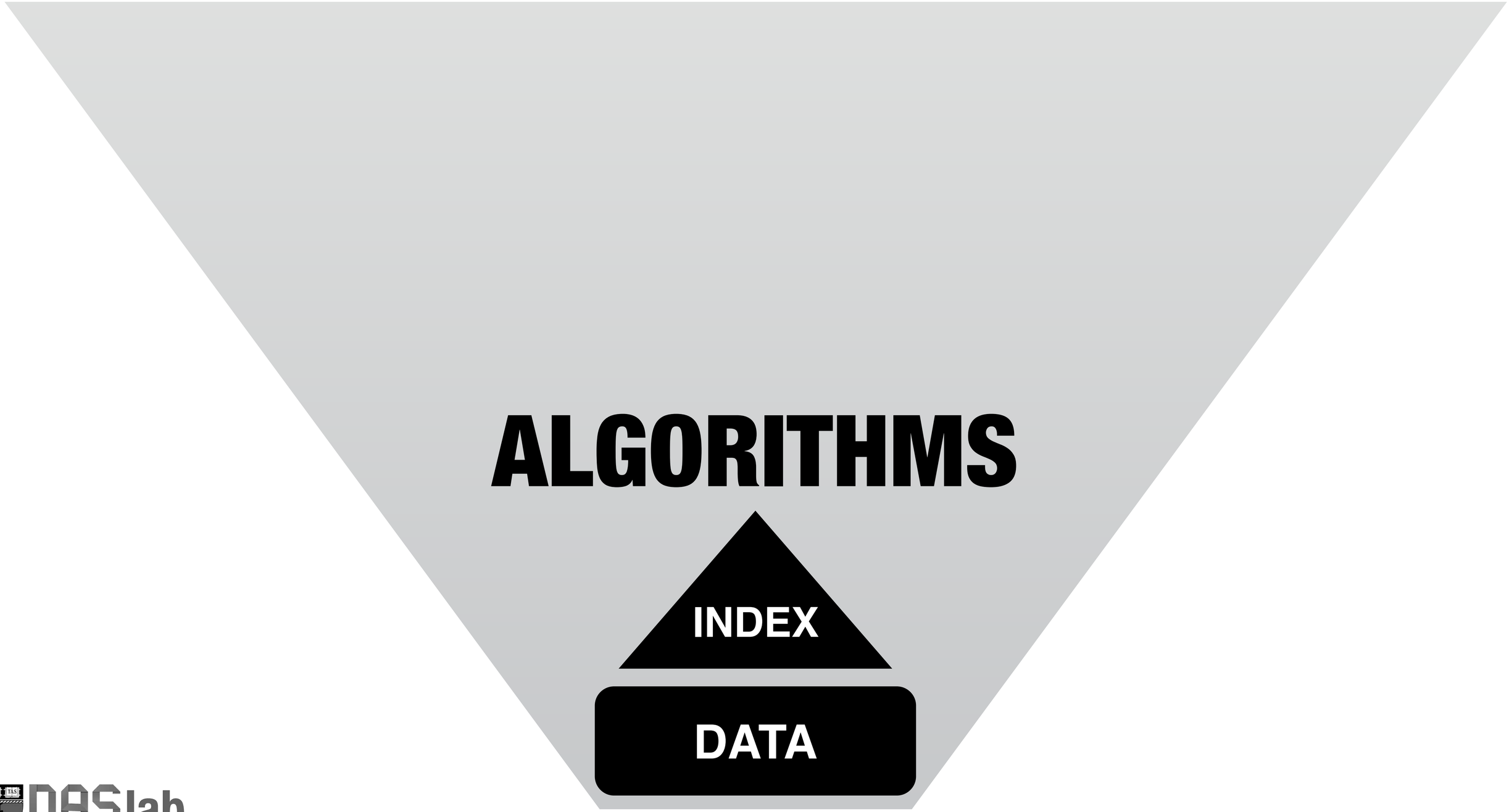
[7,4,2,6,1,3,9,10,5,8]

ordered

[1,2,3,4,5,6,7,8,9,10]

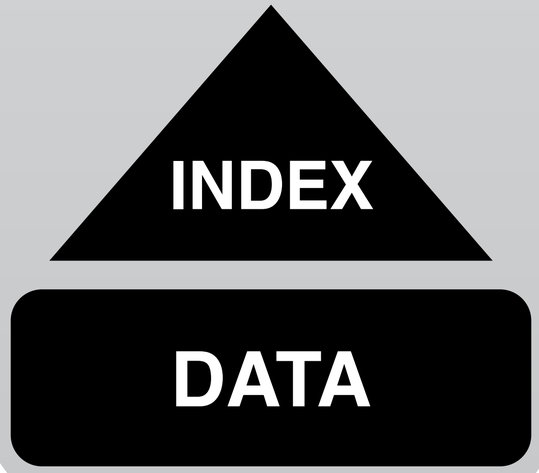
INDEX

DATA

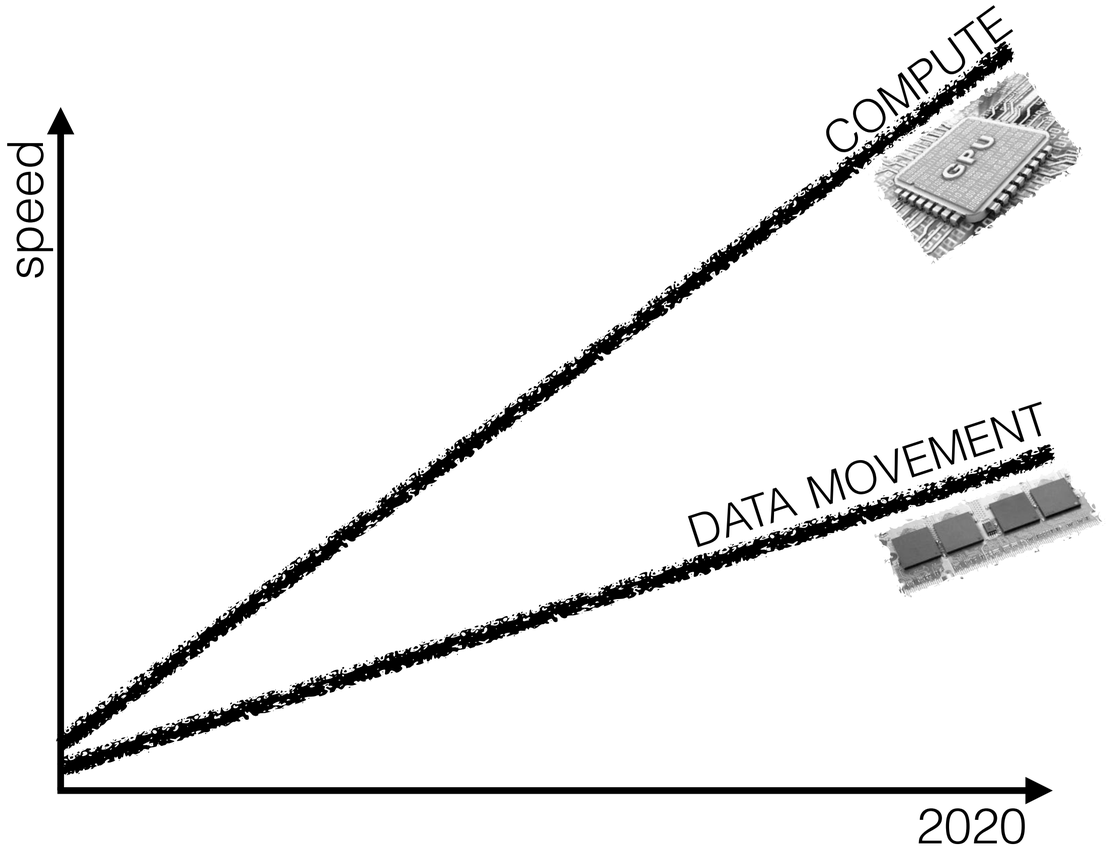

ALGORITHMS
INDEX
DATA

DATA SYSTEMS

ALGORITHMS

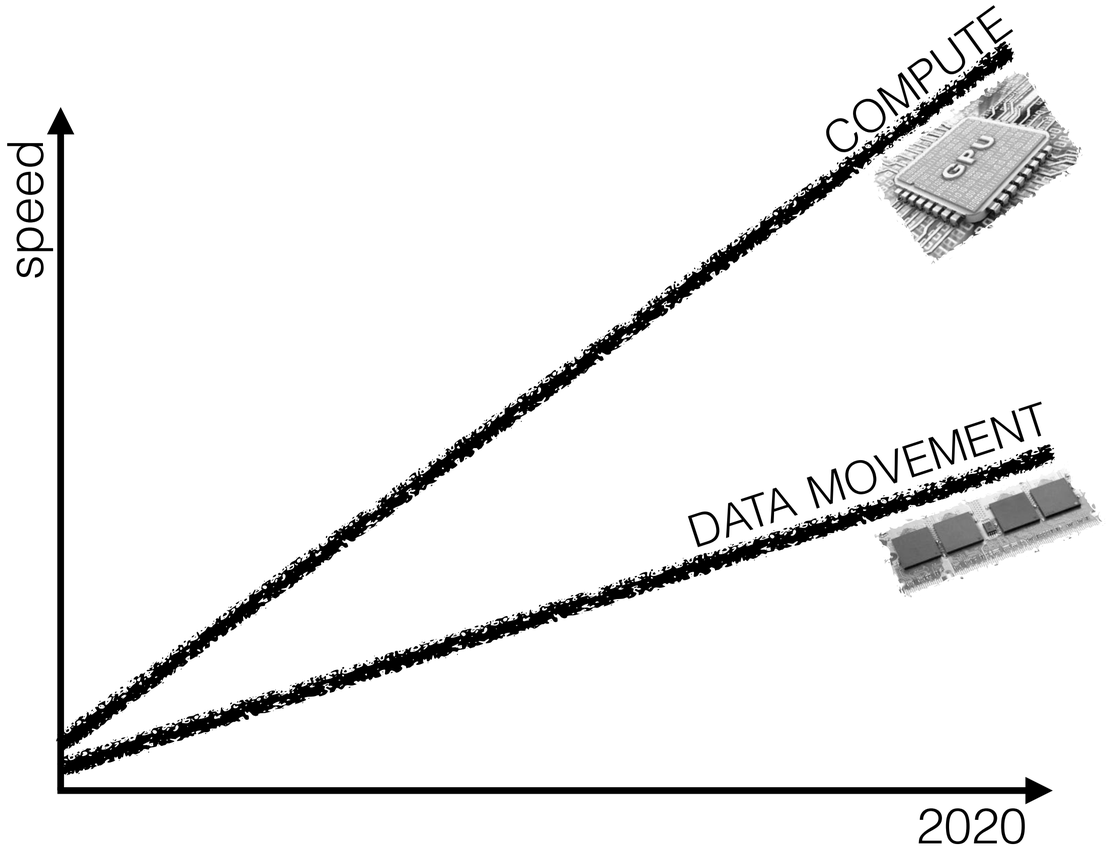


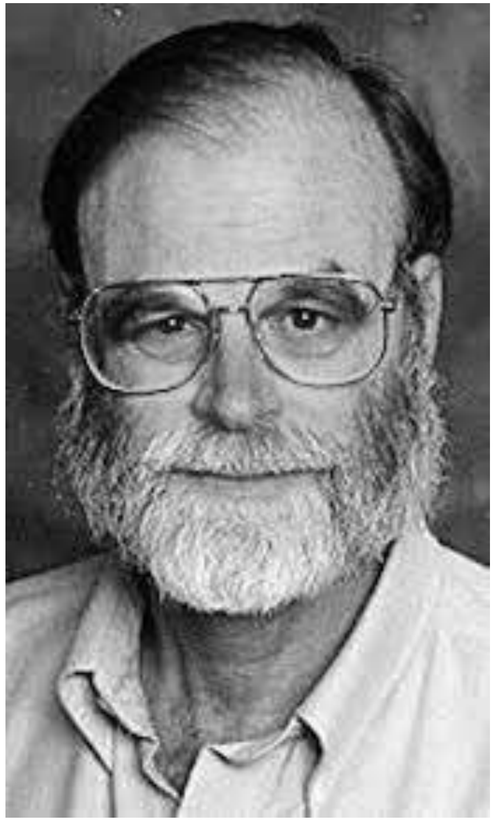
DASlab
@ Harvard SEAS



DATA STRUCTURES DEFINE PERFORMANCE

DASlab
@ Harvard SEAS





register = this room

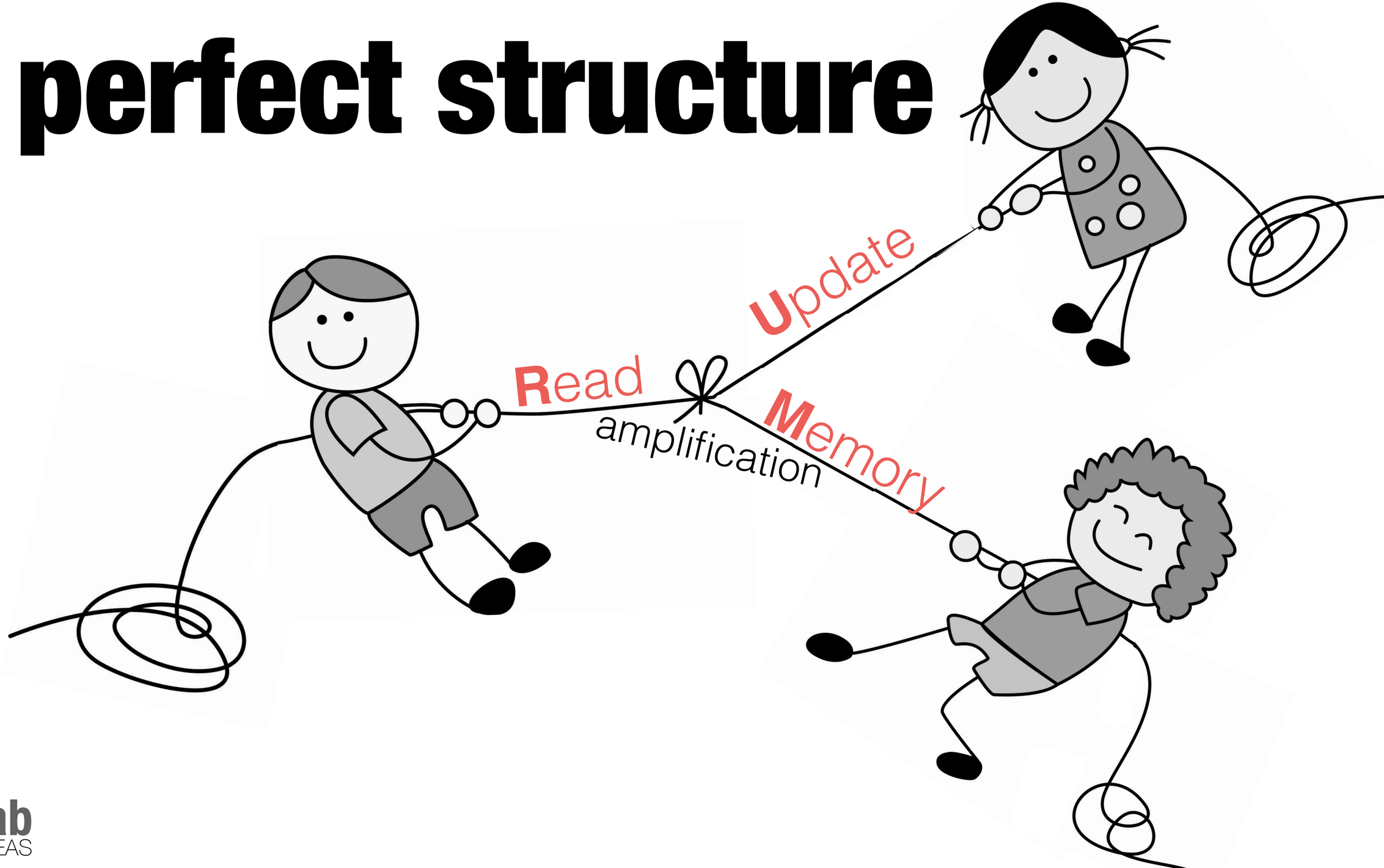
caches = this city

memory = nearby city

disk = Pluto

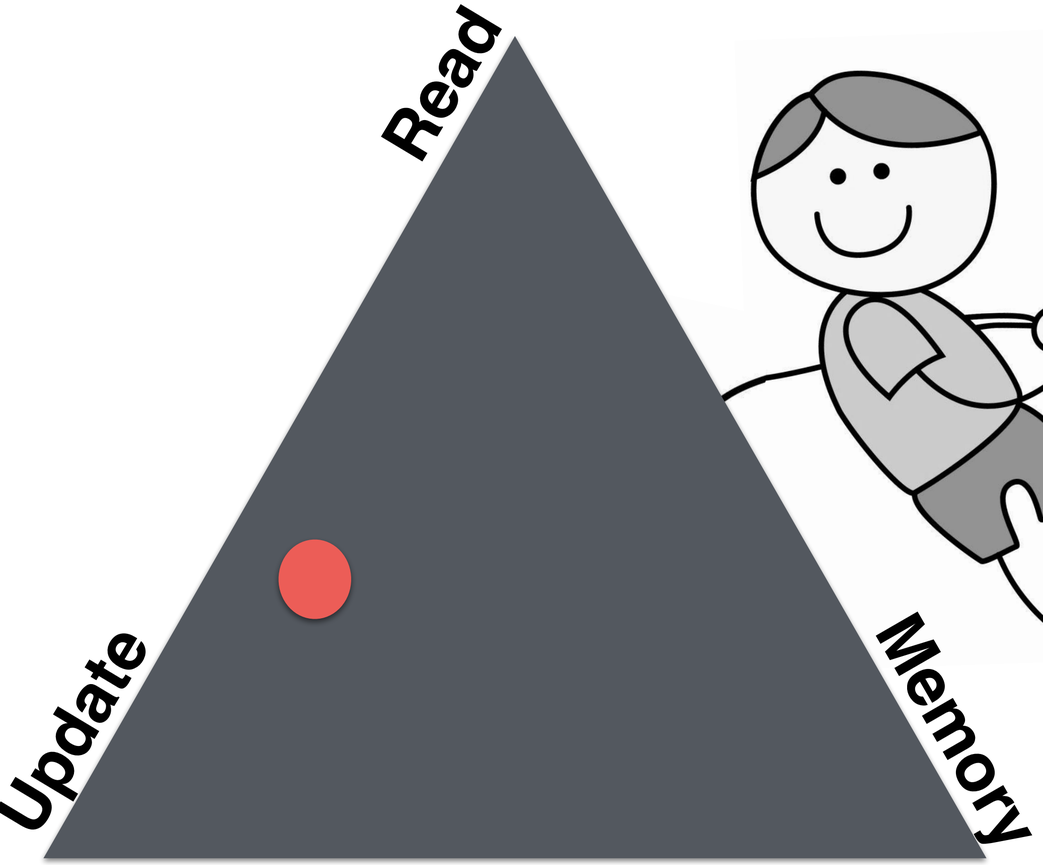
Jim Gray, Turing Award 1998

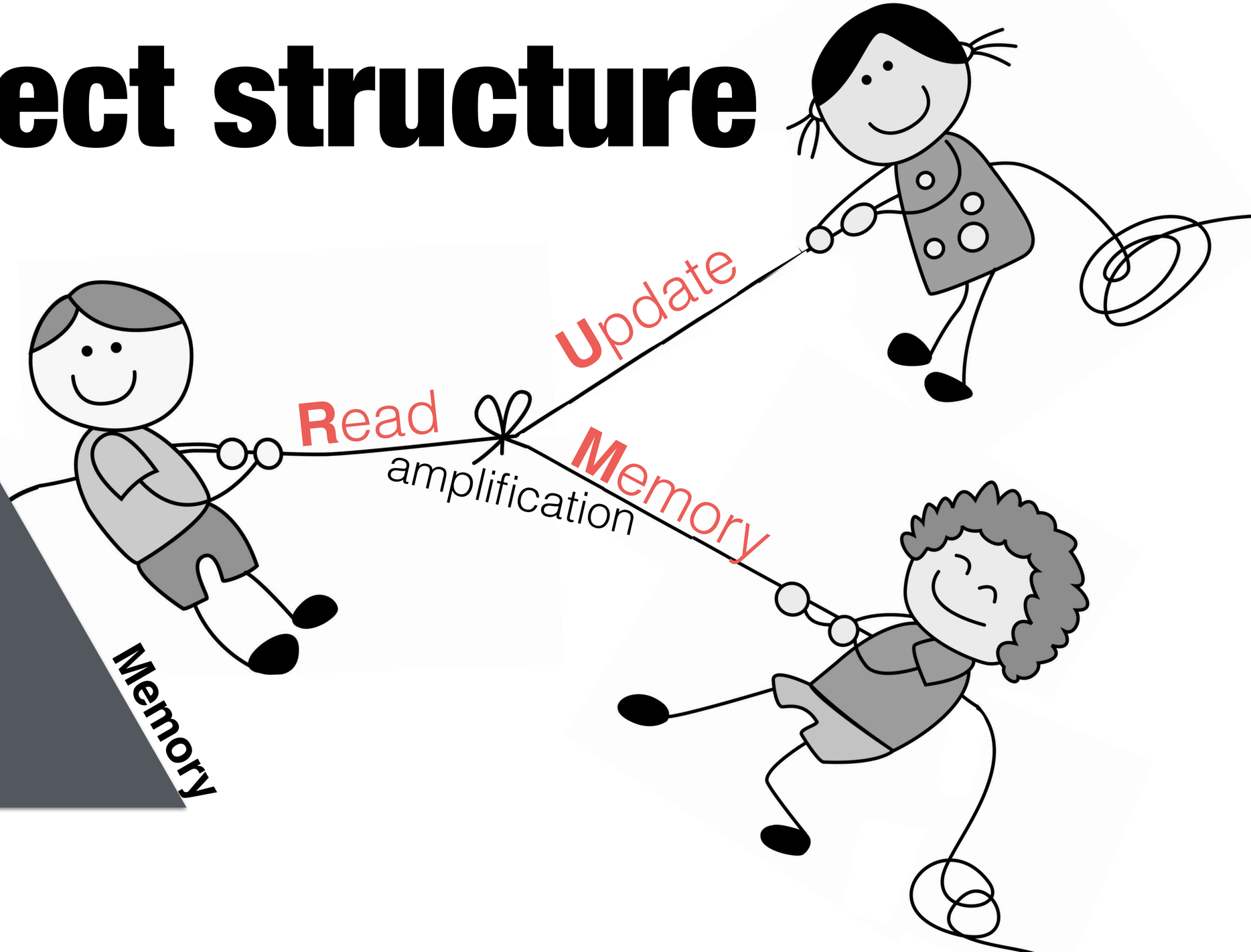
no perfect structure



DASlab
@ Harvard SEAS

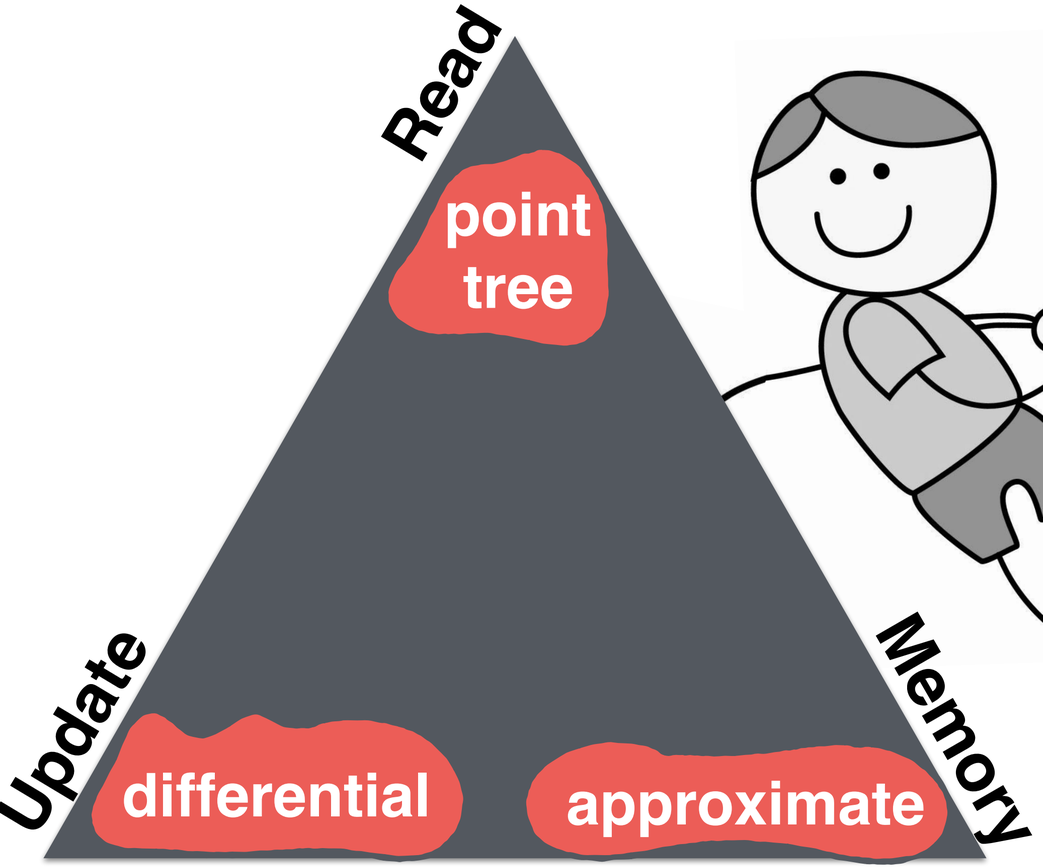
no perfect structure

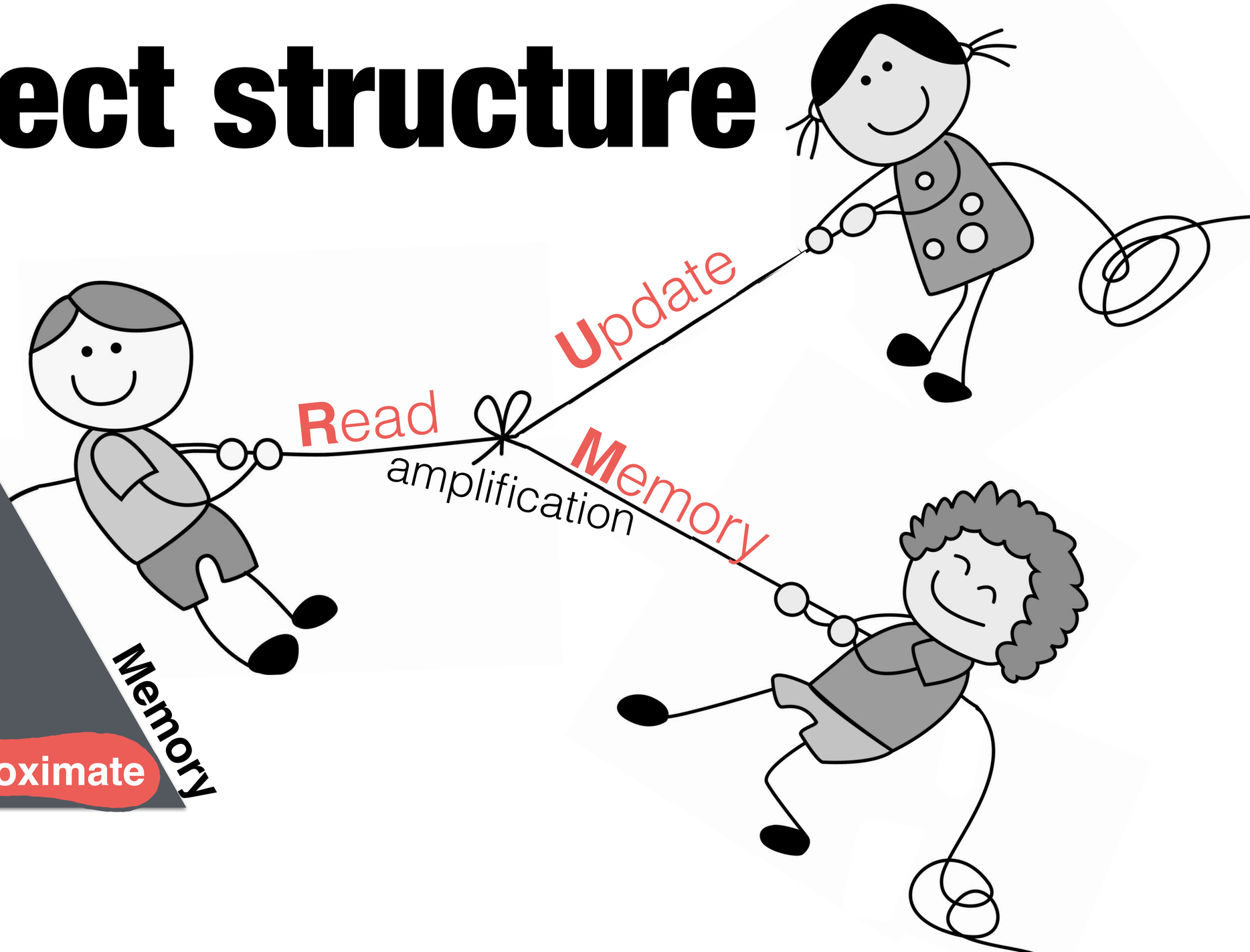




DASlab
@ Harvard SEAS

no perfect structure





DASlab
@ Harvard SEAS

no perfect structure

Array

Hash-Table

Skip-List

B-tree

Linked-List

Sorted Array

Trie

How do I make my **data system** run x times as fast?  (sql,nosql,bigdata, …)

How do I make my **data system** run x times as fast?  (sql,nosql,bigdata, …)

How do I minimize my **bill** in the **cloud**?

How do I make my **data system** run x times as fast?  (sql,nosql,bigdata, …)

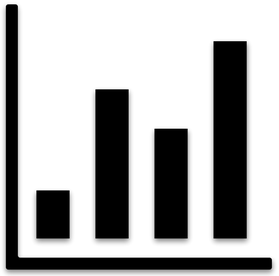
How do I minimize my **bill** in the **cloud**?

How do I extend the **lifetime** of my hardware? 

How do I make my **data system** run x times as fast?  (sql,nosql,bigdata, …)

How do I minimize my **bill** in the **cloud**?

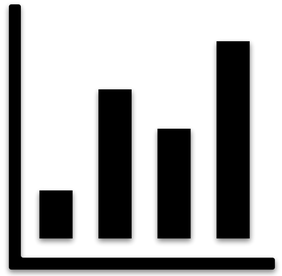
How do I extend the **lifetime** of my hardware? 

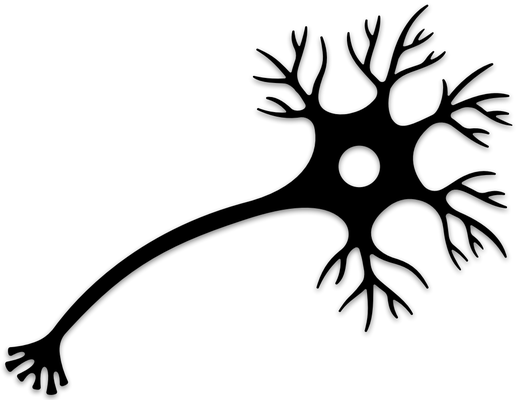
How to accelerate **statistics** computation for data science/ML? 

How do I make my **data system** run x times as fast?  (sql,nosql,bigdata, …)

How do I minimize my **bill** in the **cloud**?

How do I extend the **lifetime** of my hardware? 

How to accelerate **statistics** computation for data science/ML? 



How do I train my **neural network** x times faster?

DASlab @ Harvard SEAS

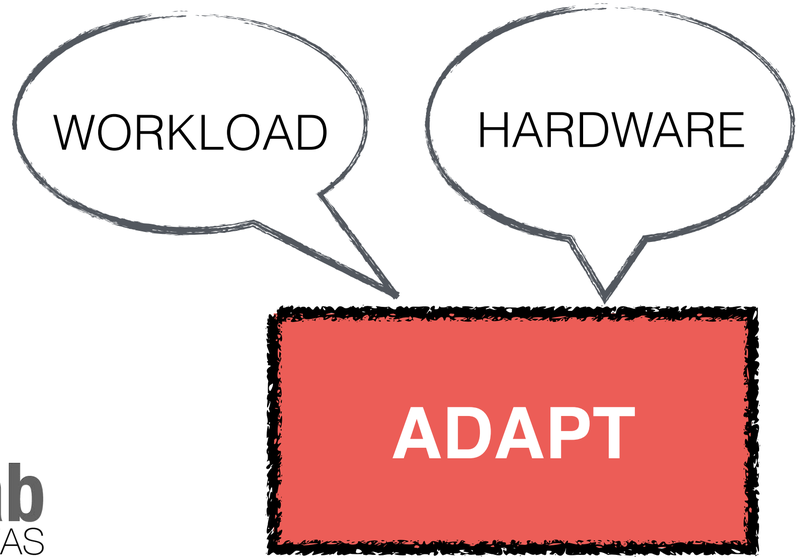
NEW APPLICATIONS

NEW APPLICATIONS

existing systems need to change too

NEW APPLICATIONS

existing systems need to change too



DASlab
@ Harvard SEAS

NEW APPLICATIONS

existing systems need to change too

WORKLOAD

HARDWARE

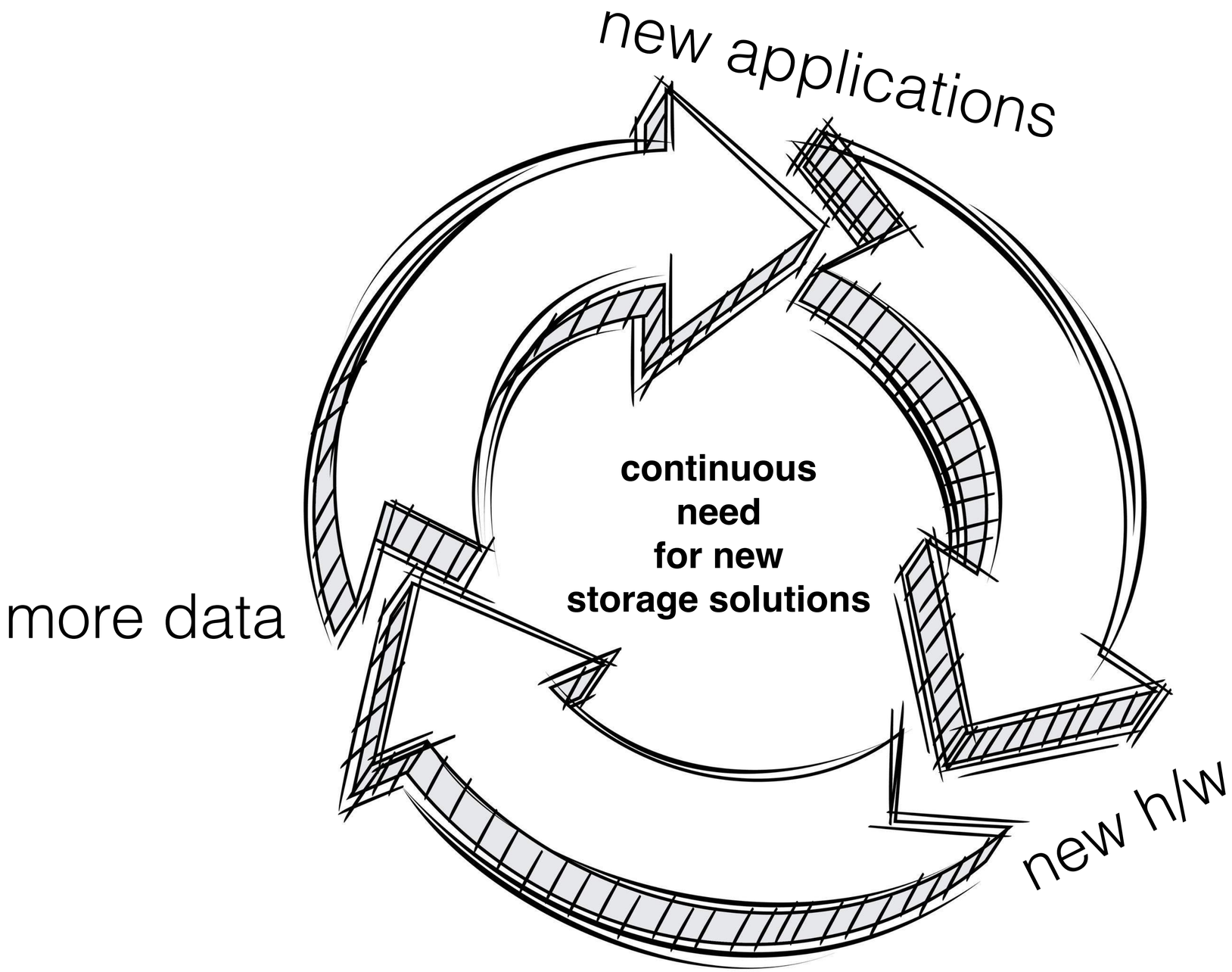
ADAPT

IMPROVE WITHIN A BUDGET

WHAT WILL BREAK MY SYSTEM?

REASON

DASlab @ Harvard SEAS


new applications
continuous
need
for new
storage solutions
more data
new h/w

learning outcome
fundamental of storage

learning outcome
fundamental of storage

software engineering **data-driven startup** **research**

learning outcome
fundamental of storage

software engineering **data-driven startup** **research**

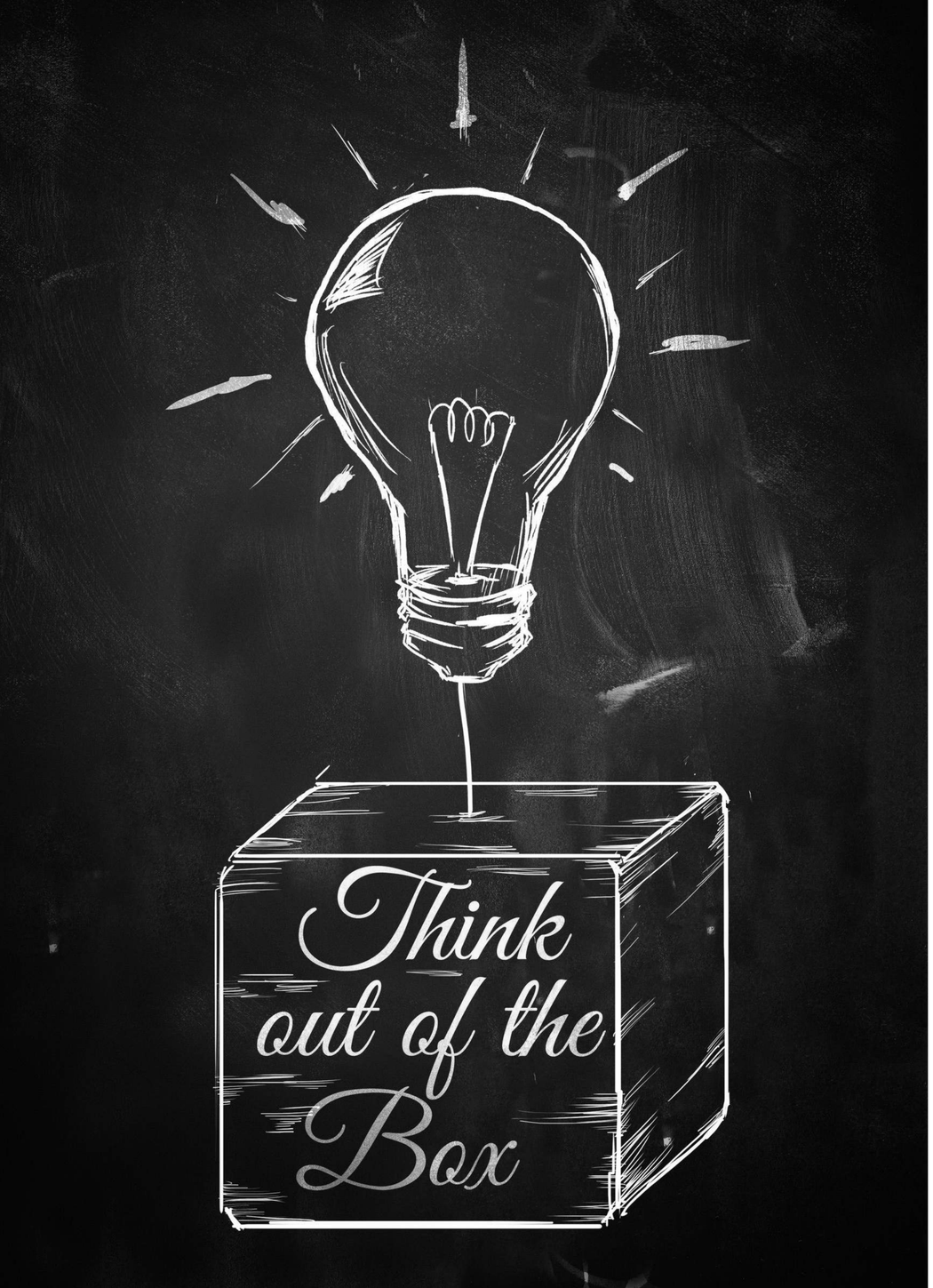
data structures, SQL, NoSQL, Big Data, Neural Networks, Statistics, Data Science

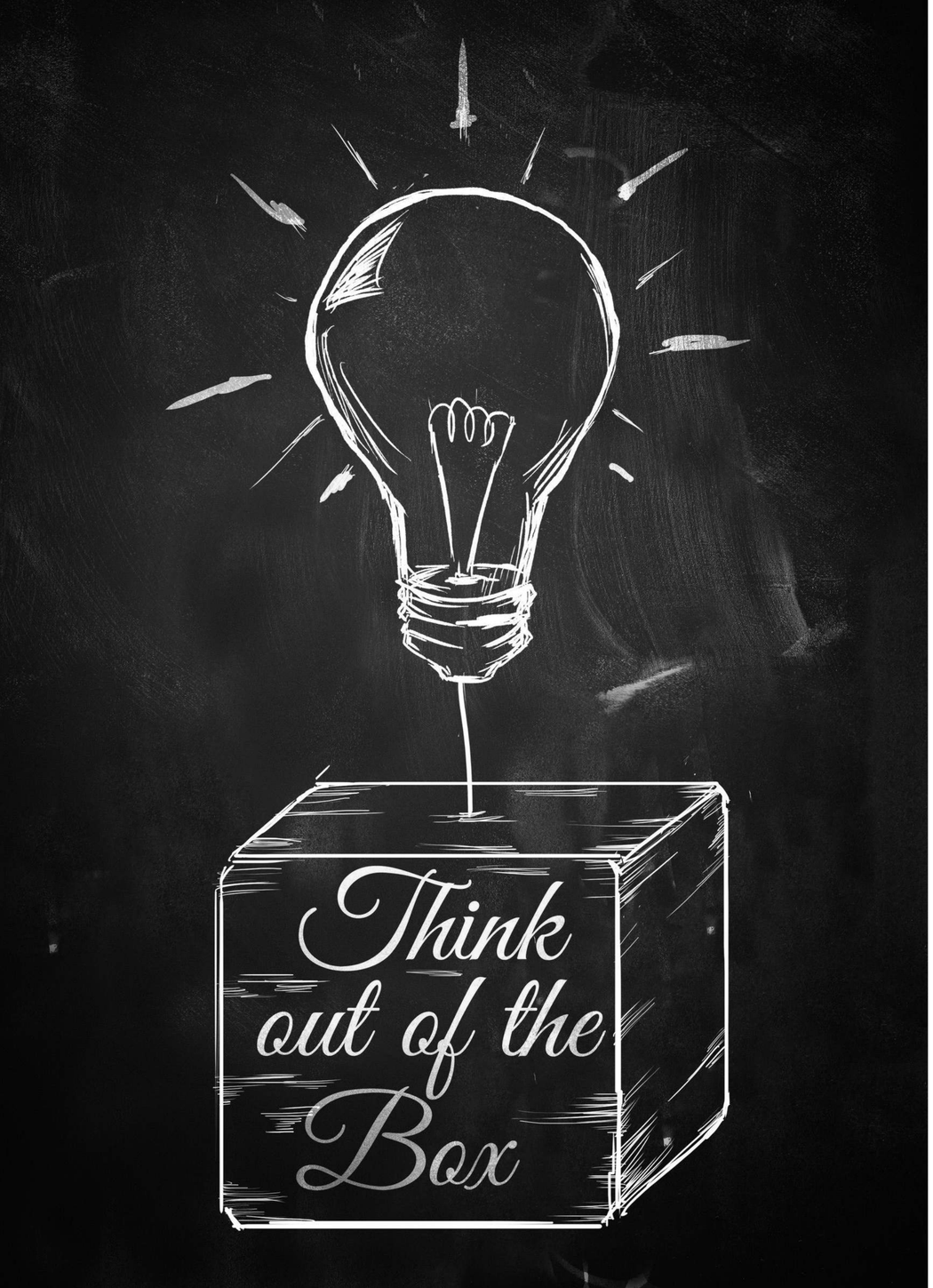
learning outcome
fundamental of storage

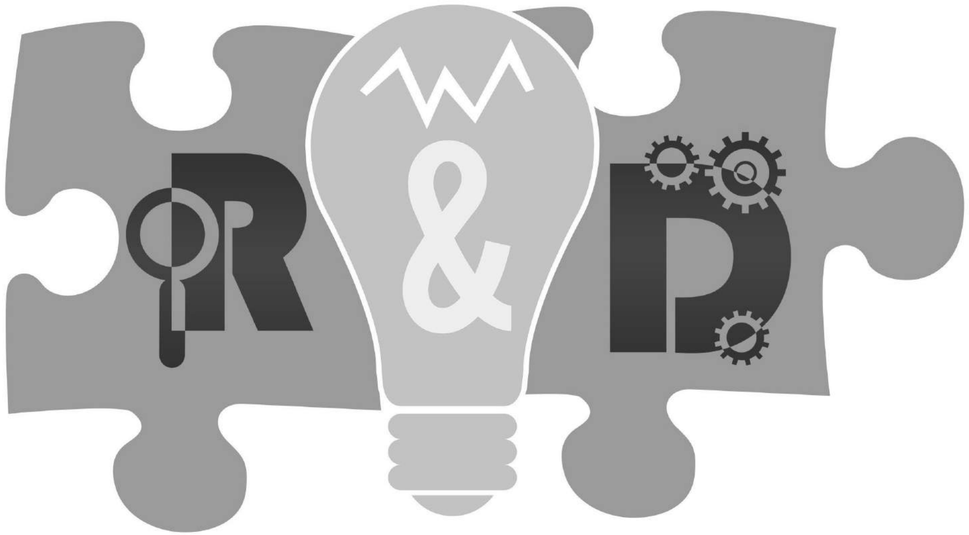
software engineering **data-driven startup** **research**

data structures, SQL, NoSQL, Big Data, Neural Networks, Statistics, Data Science

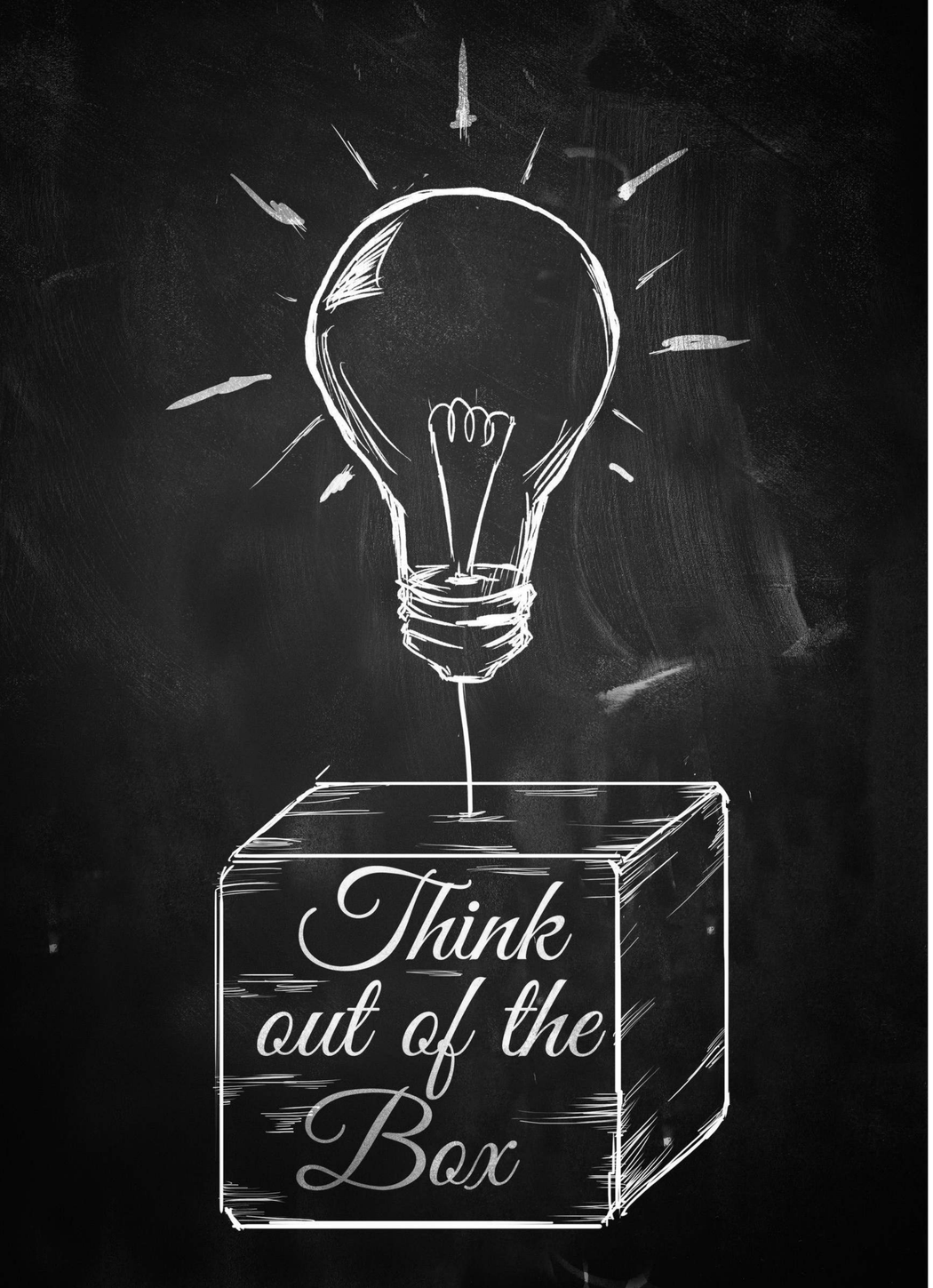
small set of principles across all fields

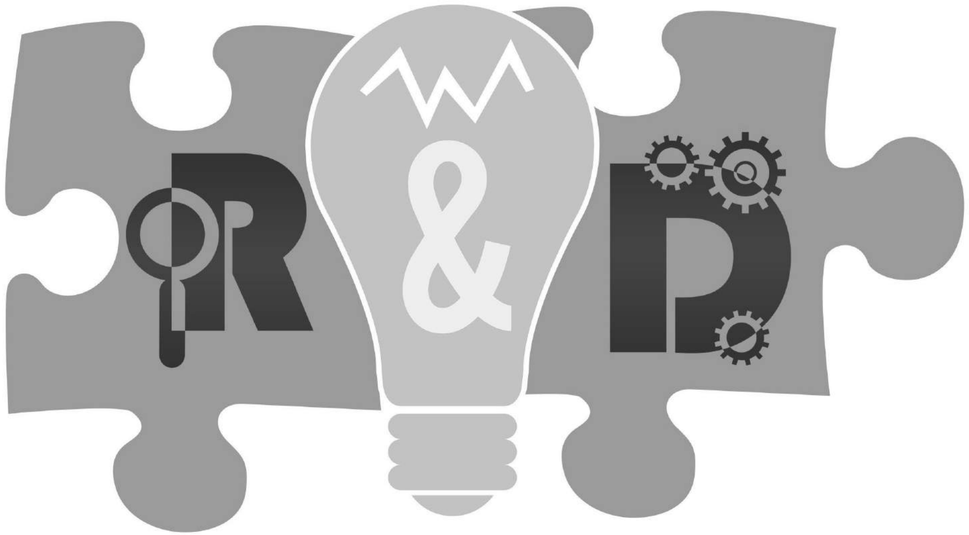

Think
out of the
Box





first 4 weeks: introduction to research problems/thinking through lectures



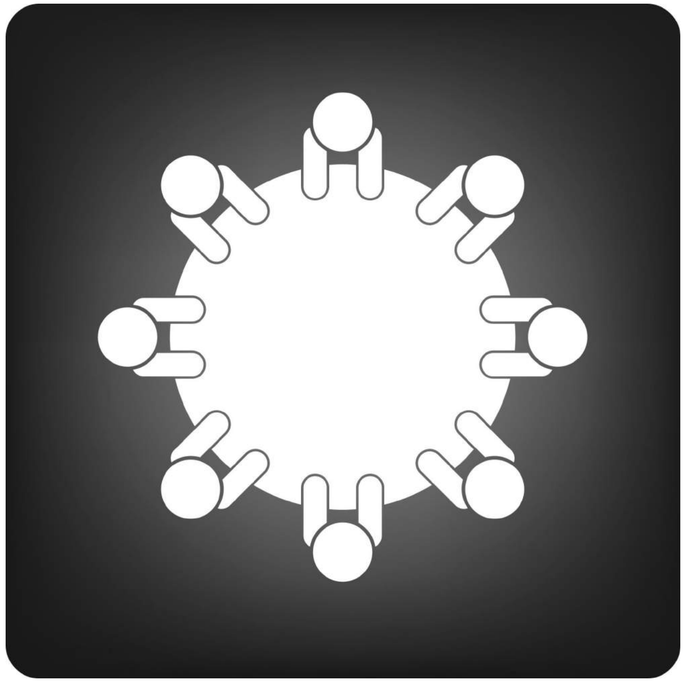


first 4 weeks: introduction to research problems/thinking through lectures

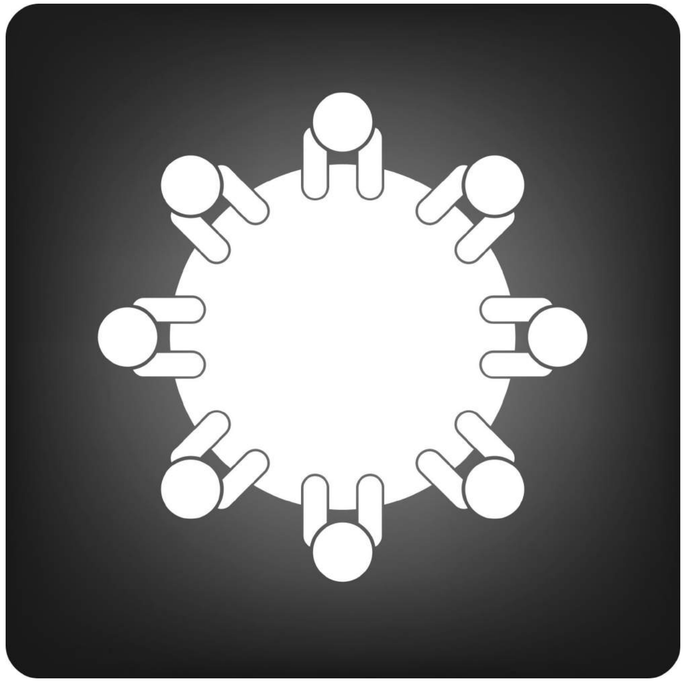
Reading research papers

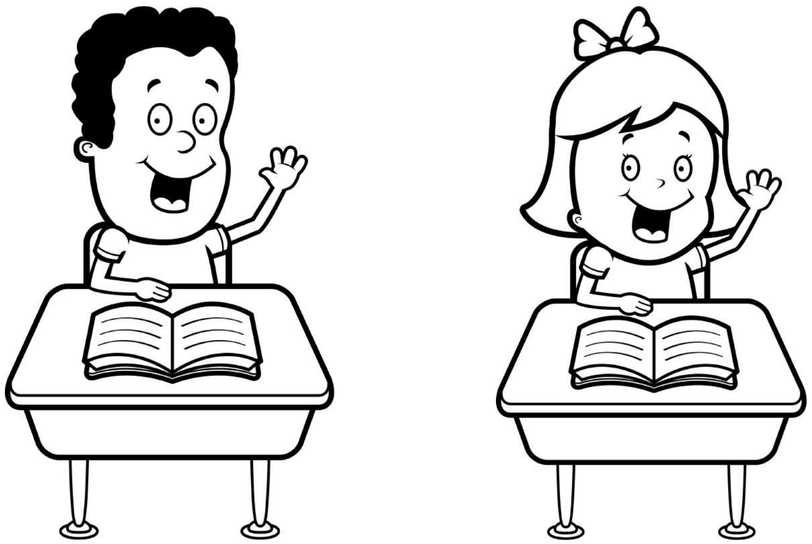
Open ended projects/research

as of week 5:
discussions/
presentations



as of week 5:
discussions/
presentations



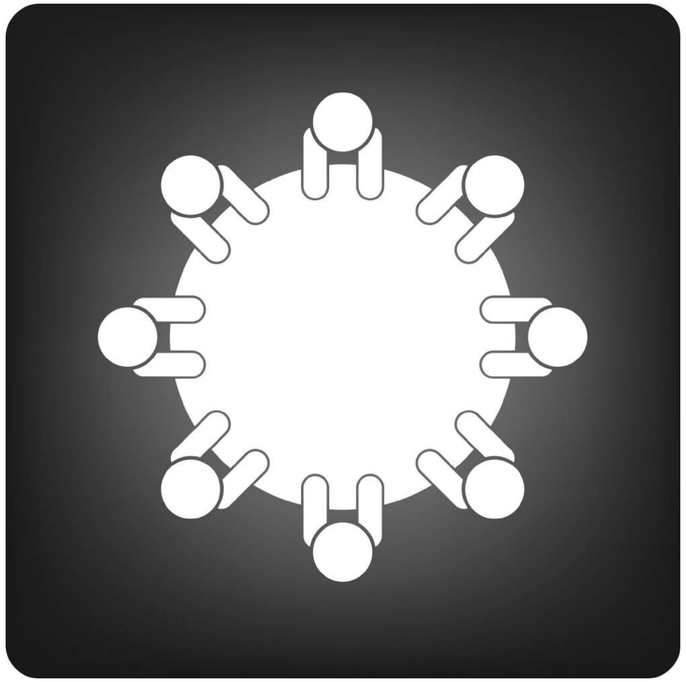


interaction: in and out of class

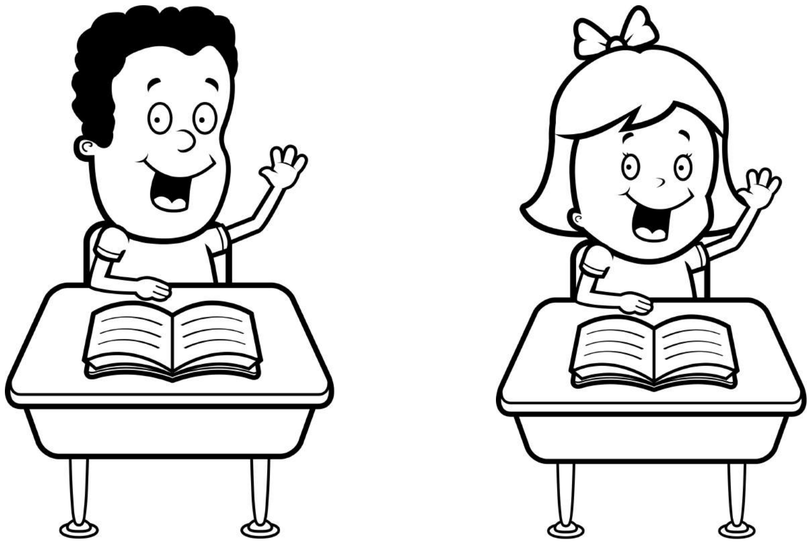
M/W/F OH/labs, Sat/Sun remote OH

DASlab
@ Harvard SEAS

as of week 5:
discussions/
presentations



There is no such thing as a wrong question/answer!!!!



interaction: in and out of class

M/W/F OH/labs, Sat/Sun remote OH

DASlab @ Harvard SEAS

Recent Research Papers

Each student:

2 reviews per week/1 presentation

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

learn to judge constructively

learn to present

learn to prepare slides

Each student:

2 reviews per week/1 presentation

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

semester project: due in the end of semester + a midway check in (early March,10%)

systems project

research project

semester project: due in the end of semester + a midway check in (early March,10%)

systems project

individual project
NoSQL, in c/c++

research project

semester project: due in the end of semester + a midway check in (early March,10%)

systems project

individual project
NoSQL, in c/c++

research project

groups of three
NoSQL, Neural Networks
Periodic Table of Data Structures

semester project: due in the end of semester + a midway check in (early March,10%)

systems project

individual project
NoSQL, in c/c++

research project

groups of three
NoSQL, Neural Networks
Periodic Table of Data Structures

ACM Special Interest Group In Data Management (SIGMOD)

Undergrad Research Competition

first prize in 2016, 2017, 2018, 2019

Adaptive Denormalization
Evolving Trees
Splaying LSM-Trees
Adaptive NoSQL

DASlab
@ Harvard SEAS

ACM Special Interest Group In Data Management (SIGMOD)

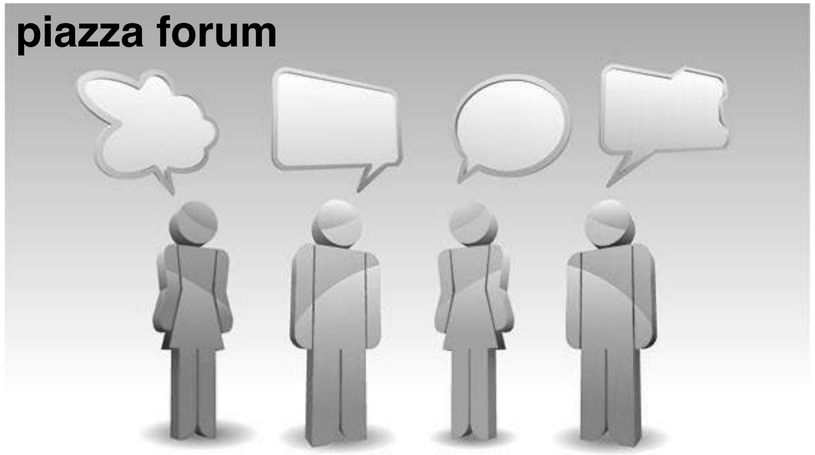
Undergrad Research Competition

first prize in 2016, 2017, 2018, 2019

Adaptive Denormalization
Evolving Trees
Splaying LSM-Trees
Adaptive NoSQL

Design continuums at CIDR 2019, two projects in SIGMOD 2020 finals

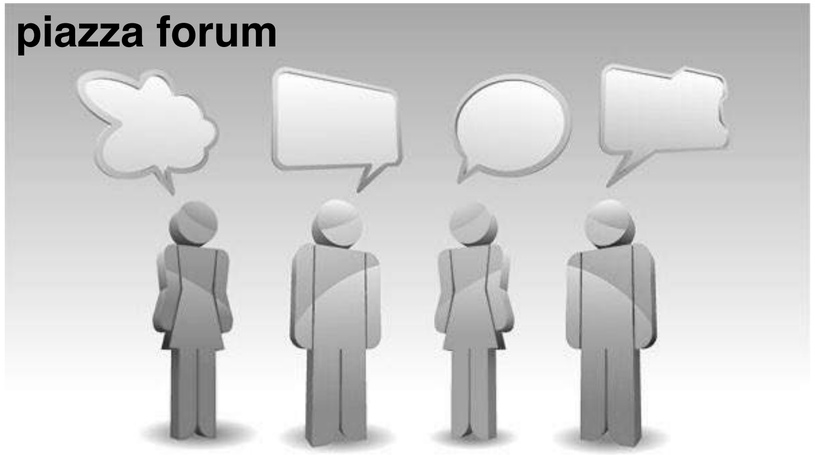
DASlab @ Harvard SEAS



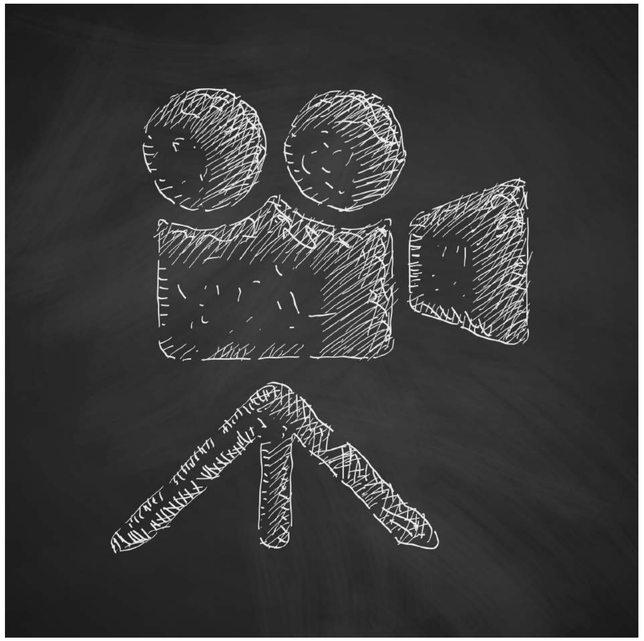
all announcements & discussions

as of week 2

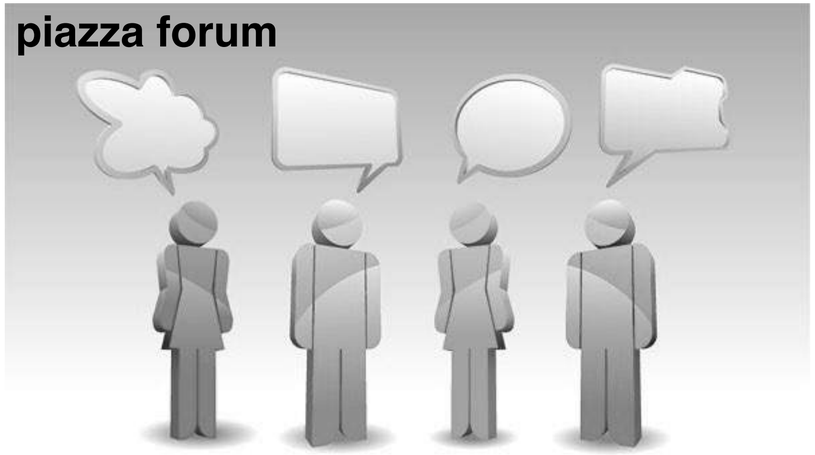
link on class website - check out usage guidelines



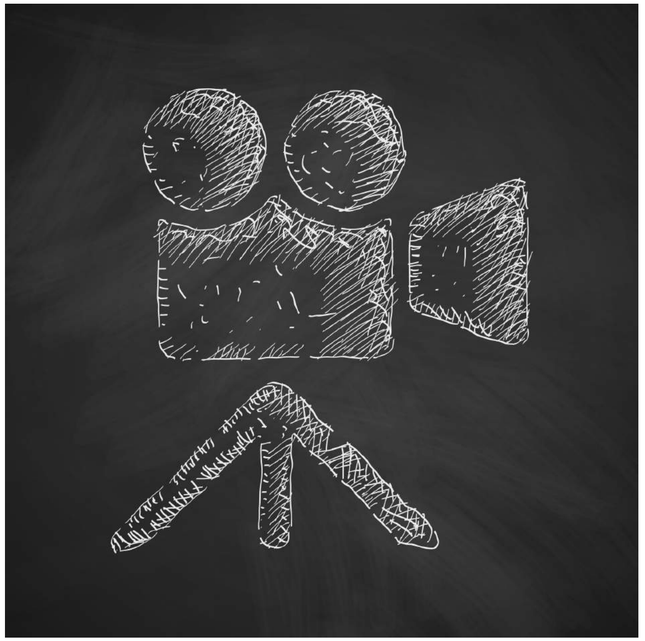
all announcements & discussions
as of week 2
link on class website - check out usage guidelines



classes are recorded
(links on class website)



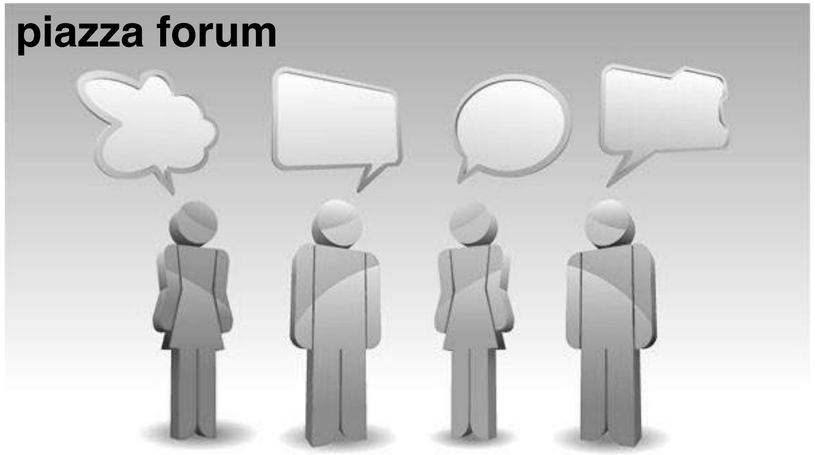
all announcements & discussions
as of week 2
link on class website - check out usage guidelines



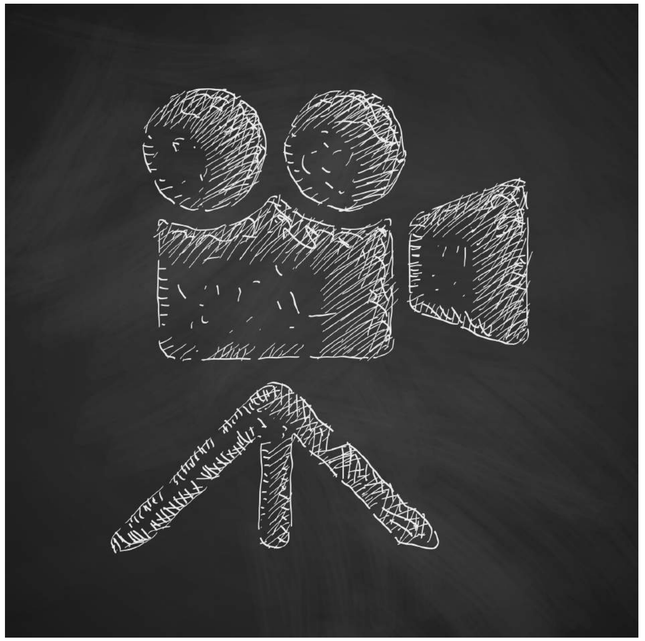
classes are recorded
(links on class website)

NO LAPTOP/PHONE POLICY
class is based on participation!

DASlab
@ Harvard SEAS



all announcements & discussions
as of week 2
link on class website - check out usage guidelines



classes are recorded
(links on class website)

Project: 40%
Midway Check-in:10%
Discussion: 20%
Presentation: 15%
Reviews: 15%

NO LAPTOP/PHONE POLICY
class is based on participation!

DASlab
@ Harvard SEAS

Get familiar with the very basics of traditional database architectures:
Architecture of a Database System. By J. Hellerstein, M. Stonebraker and J. Hamilton. Foundations and Trends in Databases, 2007

Get familiar with very basics of modern database architectures:
The Design and Implementation of Modern Column-store Database Systems. By D. Abadi, P. Boncz, S. Harizopoulos, S. Idreos, S. Madden. Foundations and Trends in Databases, 2013

Get familiar with the very basics of modern large scale systems:
Massively Parallel Databases and MapReduce Systems. By Shivnath Babu and Herodotos Herodotou. Foundations and Trends in Databases, 2013

Check out: syllabus, preparation readings, project 0, systems project, online sections

http://daslab.seas.harvard.edu/classes/cs265/

DASlab @ Harvard SEAS

Teaching Fellows:

Off class discussions are key! question on readings, ideas, help with code/analysis

DASlab
@ Harvard SEAS

Prerequisites

knowledge of algorithms, data structures, hardware, systems

Prerequisites

knowledge of algorithms, data structures, hardware, systems

Research track: open to CS165 students

Systems track allows taking the class without all prerequisites

Prerequisites

knowledge of algorithms, data structures, hardware, systems

Research track: open to CS165 students

Systems track allows taking the class without all prerequisites

(165/265 will not be offered in fall 2020/spring2021)

questions on logistics?

Next few classes:

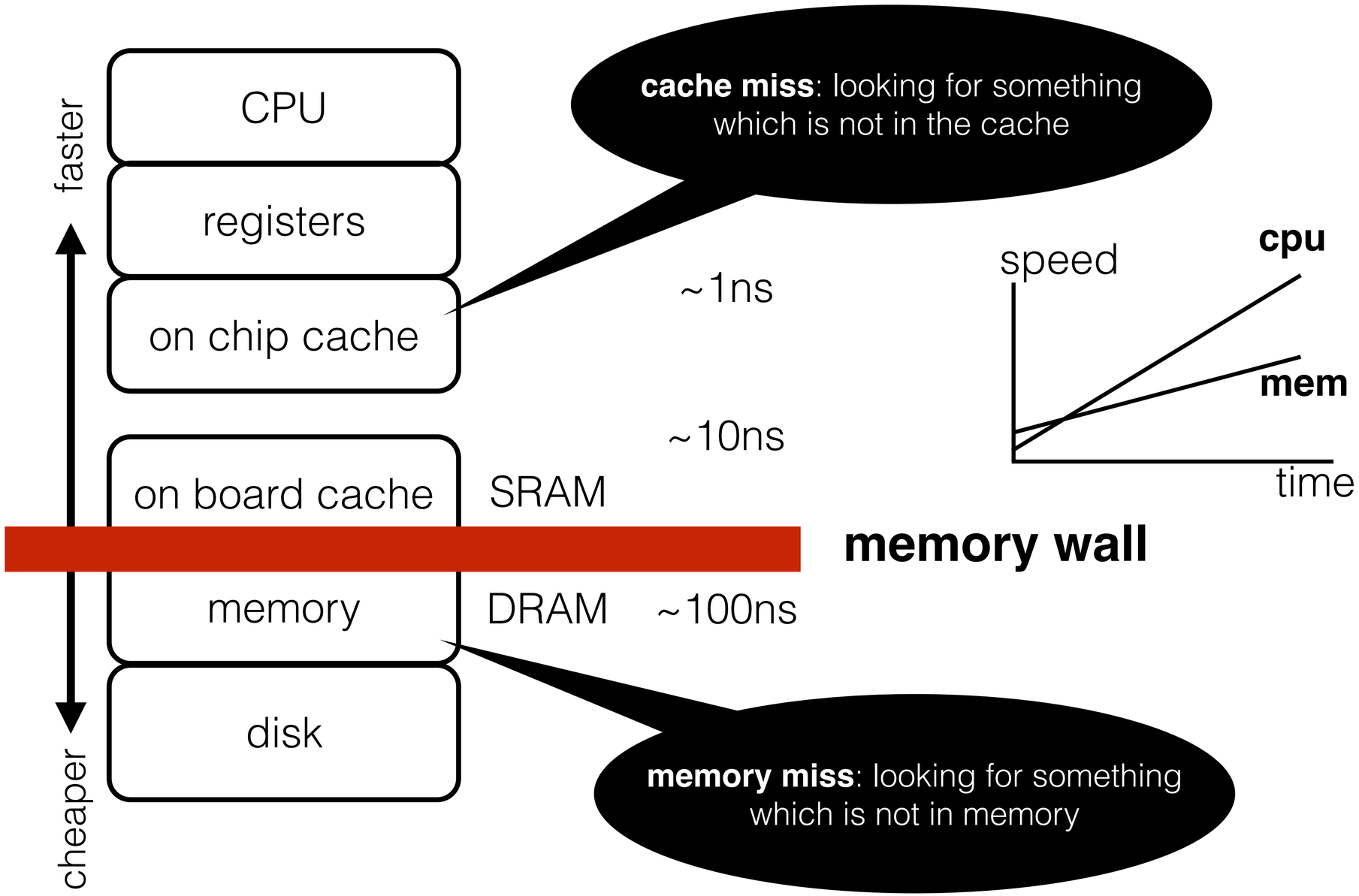
BASICS of storage

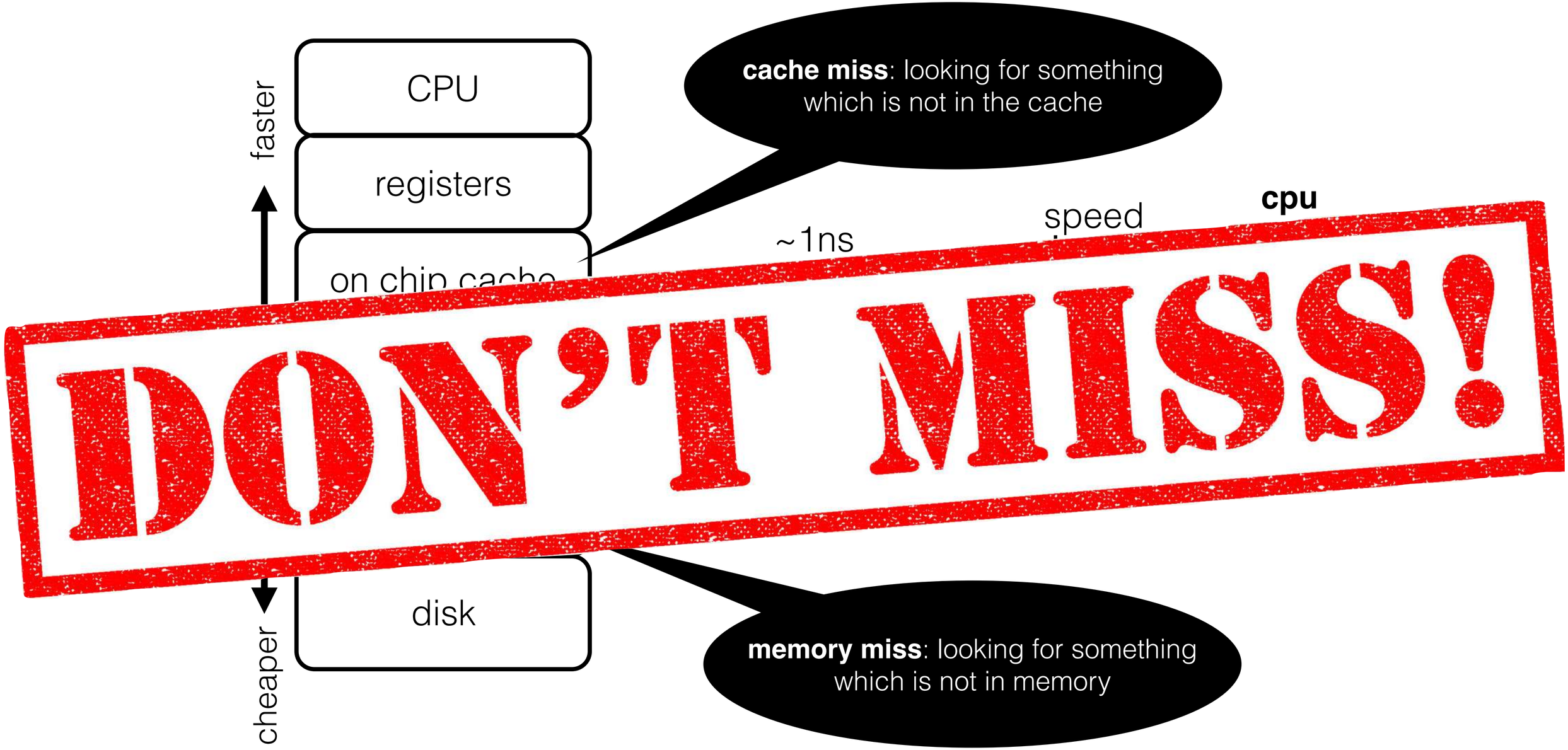
Intro to RESEACH topics

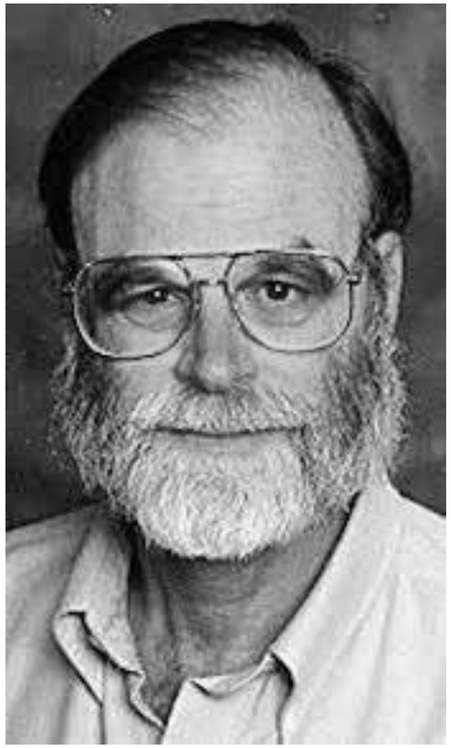
Discussion phase/presentation as of week 5

periodic table of data structures

classes of primitives \ classes of designs	B-trees & Variants	Tries & Variants	LSM-Trees & Variants	Differential Files	Membership Tests	Zone maps & Variants	Bitmaps & Variants	Hashing	Base Data & Columns	
Partitioning	DONE	DONE	DONE					DONE	DONE	RUM
Logarithmic Design	DONE	DONE	DONE							RUM
Fractional Cascading	DONE		DONE	DONE						RUM
Log-Structured	DONE		DONE	DONE						RUM
Buffering	DONE			DONE			DONE			RUM
Differential Updates	DONE			DONE						RUM
Sparse Indexing	DONE				DONE	DONE				RUM
Adaptivity	DONE								DONE	

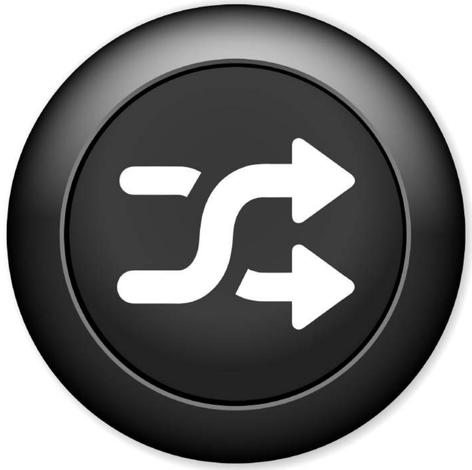

faster
cheaper
CPU
registers
on chip cache
on board cache
memory
disk
~1ns
~10ns
SRAM
DRAM
~100ns
memory wall
cache miss: looking for something which is not in the cache
memory miss: looking for something which is not in memory
speed
time
cpu
mem

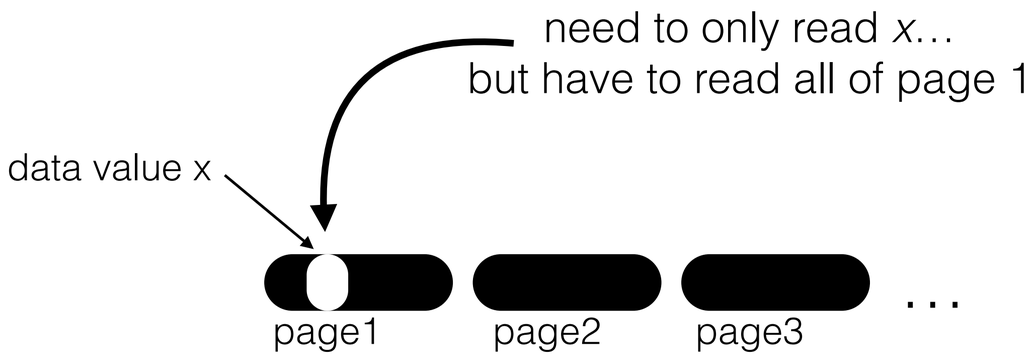

CPU
registers
on chip cache
disk
faster
cheaper
~1ns
speed
cpu
cache miss: looking for something which is not in the cache
memory miss: looking for something which is not in memory
DON'T MISS!

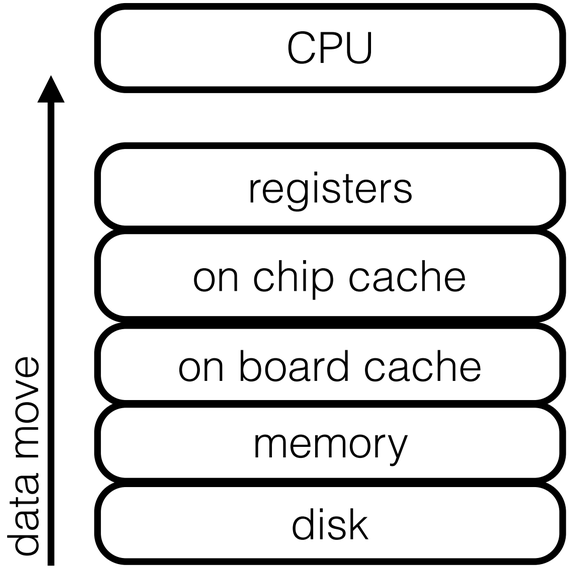


Jim Gray, IBM, Tandem, DEC, Microsoft
ACM Turing award
ACM SIGMOD Edgar F. Codd Innovations award

registers	**my head** ~0
2x on chip cache	**this room** 1 min
10x on board cache	**this building** 10 min
100x memory	**New York** 1.5 hours
100Kx disk	**Pluto** 2 years




need to only read *x*…
but have to read all of page 1
data value x
page1
page2
page3
…


CPU
registers
on chip cache
on board cache
memory
disk
data move

query x<5

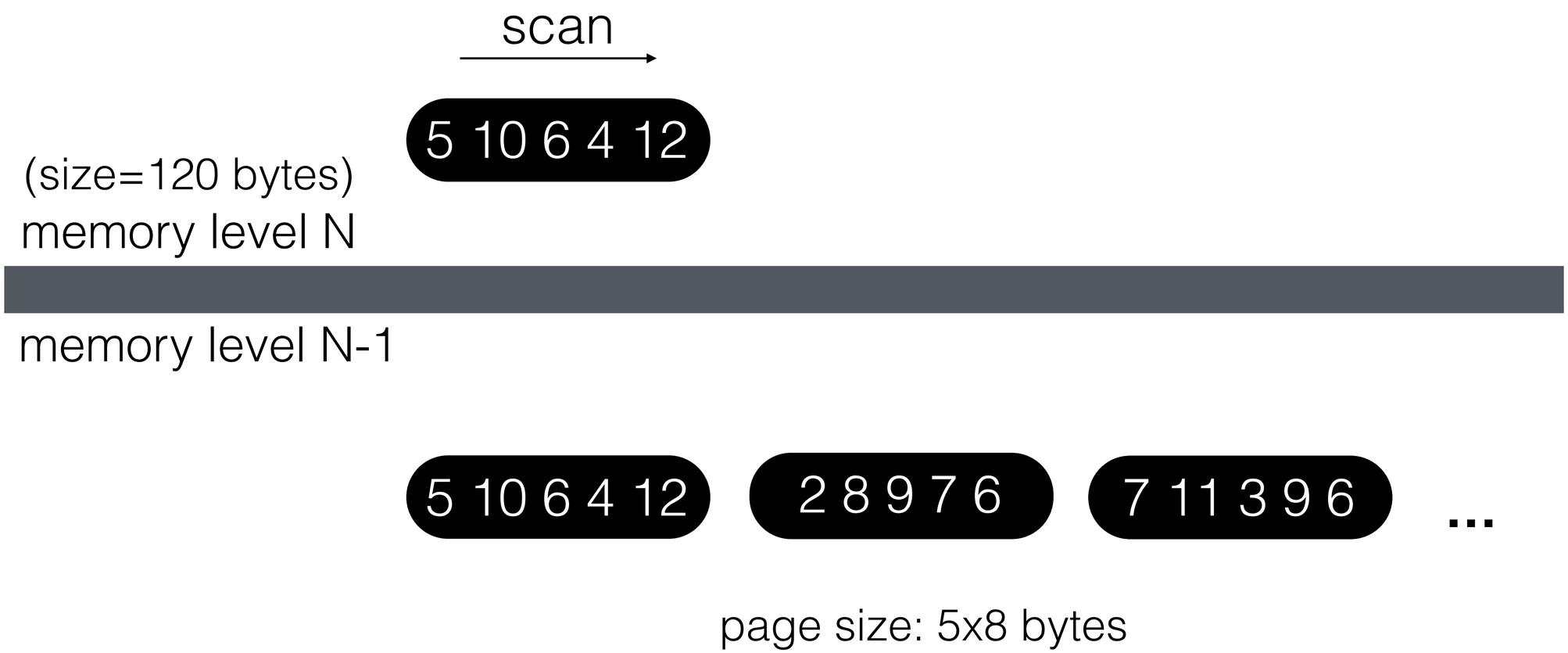
(size=120 bytes)
memory level N

memory level N-1

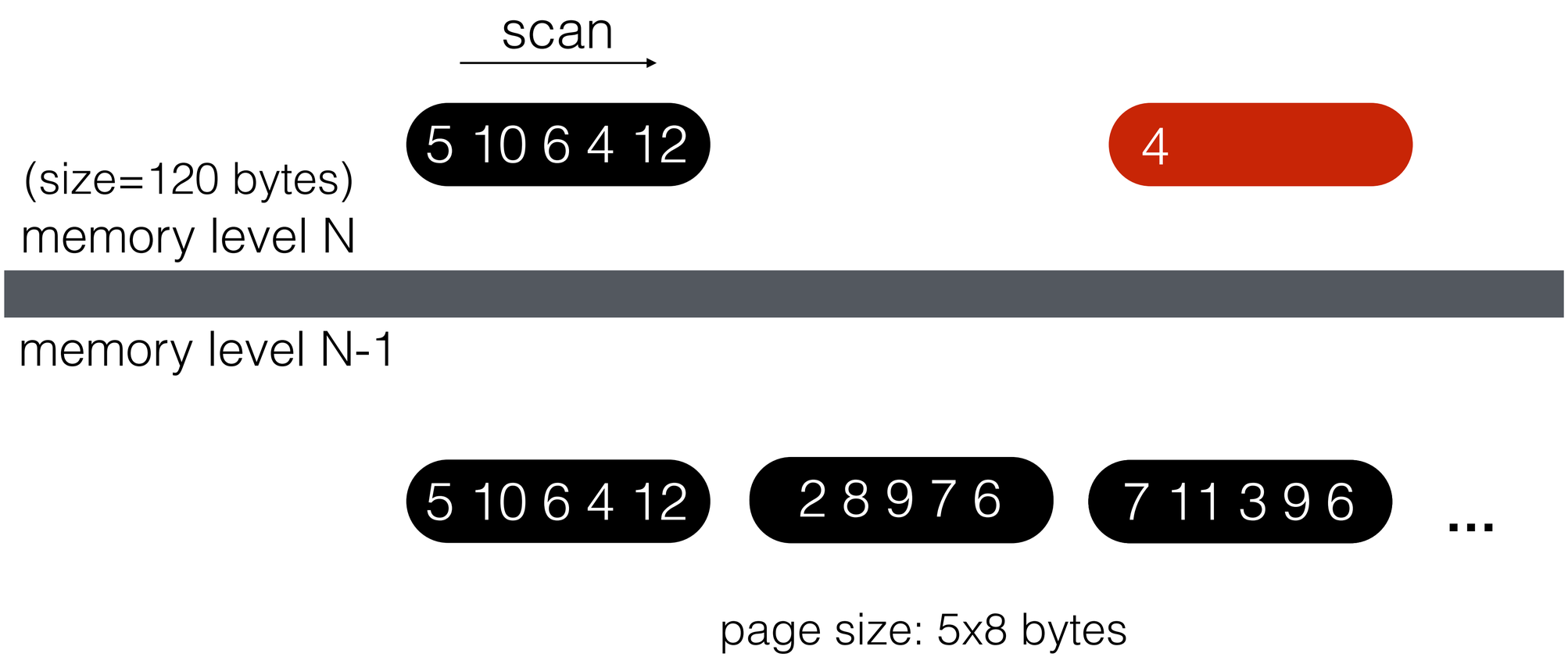
5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

query x<5


scan
5 10 6 4 12
(size=120 bytes)
memory level N
memory level N-1
5 10 6 4 12
2 8 9 7 6
7 11 3 9 6
...
page size: 5x8 bytes

query x<5


scan
5 10 6 4 12
4
(size=120 bytes)
memory level N
memory level N-1
5 10 6 4 12
2 8 9 7 6
7 11 3 9 6
...
page size: 5x8 bytes

40 bytes

query x<5

scan →

5 10 6 4 12

4

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12

2 8 9 7 6

7 11 3 9 6

...

page size: 5x8 bytes

DASlab
@ Harvard SEAS

40 bytes

query x<5

scan → scan →

5 10 6 4 12 | 2 8 9 7 6 | 4

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

40 bytes

query x<5

scan → scan →

5 10 6 4 12 | 2 8 9 7 6 | 4 2

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

DASlab @ Harvard SEAS

80 bytes

query x<5

scan → scan →

5 10 6 4 12 | 2 8 9 7 6 | 4 2

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

DASlab @ Harvard SEAS

80 bytes

query x<5

2 8 9 7 6

4 2

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12

2 8 9 7 6

7 11 3 9 6

...

page size: 5x8 bytes

DASlab
@ Harvard SEAS

80 bytes

query x<5

scan →

7 11 3 9 6

2 8 9 7 6

4 2

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12

2 8 9 7 6

7 11 3 9 6

...

page size: 5x8 bytes

80 bytes

query x<5

scan →

7 11 3 9 6 | 2 8 9 7 6 | 4 2 3

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

DASlab @ Harvard SEAS

120 bytes

query x<5

scan →

7 11 3 9 6

2 8 9 7 6

4 2 3

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12

2 8 9 7 6

7 11 3 9 6

...

page size: 5x8 bytes

an oracle gives us the positions

query x<5

(size=120 bytes)
memory level N

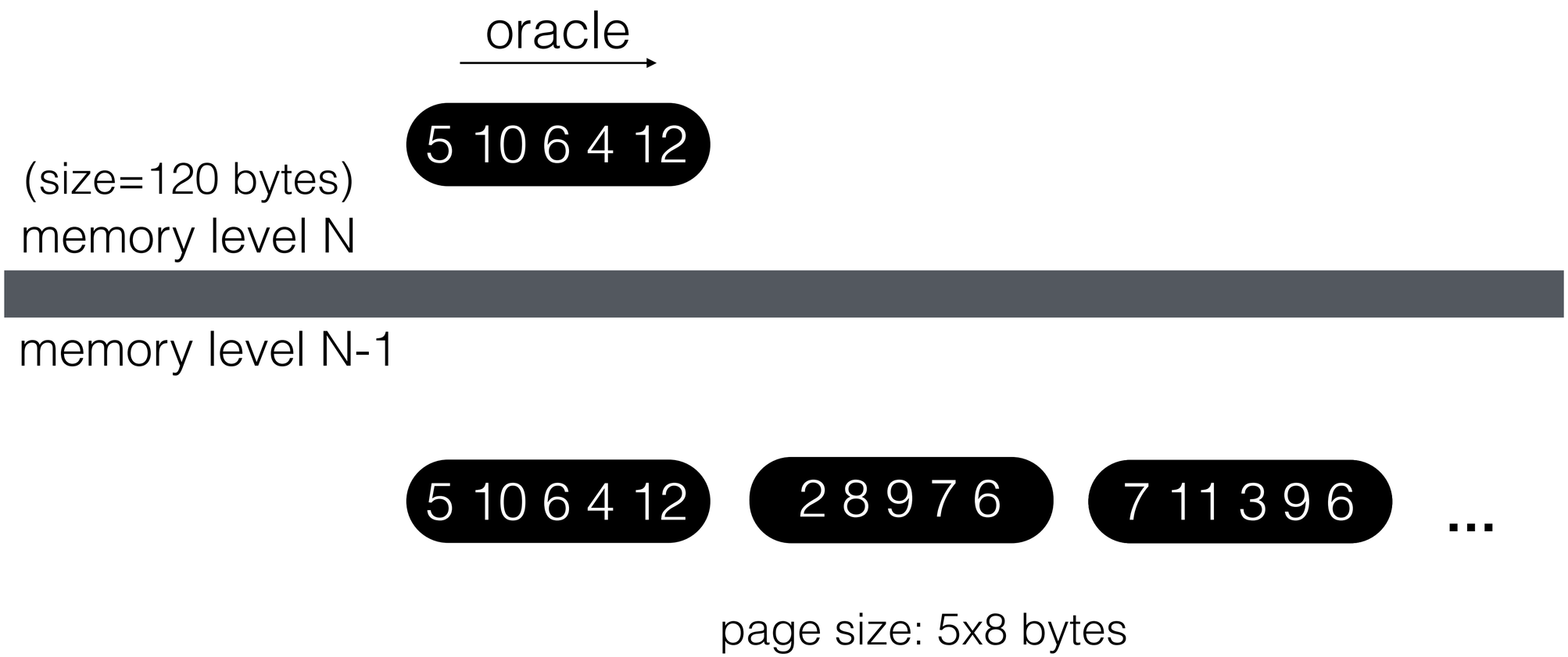
memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

an oracle gives us the positions

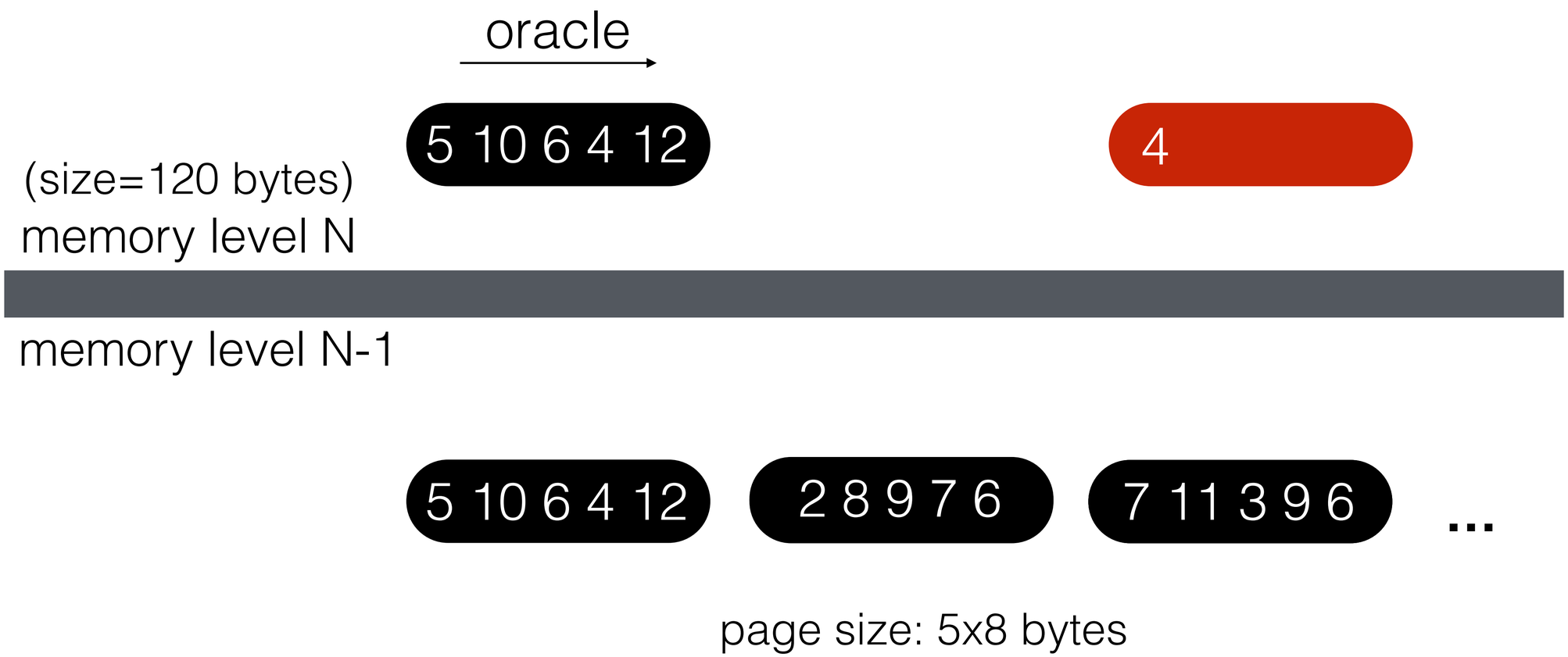
query x<5



page size: 5x8 bytes

an oracle gives us the positions

query x<5



an oracle gives us the positions

40 bytes

query x<5

oracle →

5 10 6 4 12

4

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

an oracle gives us the positions

40 bytes

query x<5

oracle → oracle →

5 10 6 4 12 | 2 8 9 7 6 | 4

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

DASlab @ Harvard SEAS

an oracle gives us the positions

40 bytes

query x<5

oracle → oracle →

5 10 6 4 12 | 2 8 9 7 6 | 4 2

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

DASlab @ Harvard SEAS

an oracle gives us the positions

80 bytes

query x<5

oracle → oracle →

5 10 6 4 12 | 2 8 9 7 6 | 4 2

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

DASlab @ Harvard SEAS

an oracle gives us the positions

80 bytes

query x<5

2 8 9 7 6

4 2

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12

2 8 9 7 6

7 11 3 9 6

...

page size: 5x8 bytes

DASlab
@ Harvard SEAS

an oracle gives us the positions

80 bytes

query x<5

oracle →

7 11 3 9 6

2 8 9 7 6

4 2

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12

2 8 9 7 6

7 11 3 9 6

...

page size: 5x8 bytes

an oracle gives us the positions

80 bytes

query x<5

oracle →

7 11 3 9 6 | 2 8 9 7 6 | 4 2 3

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

page size: 5x8 bytes

DASlab @ Harvard SEAS

an oracle gives us the positions

120 bytes

query x<5

oracle →

7 11 3 9 6

2 8 9 7 6

4 2 3

(size=120 bytes)
memory level N

memory level N-1

5 10 6 4 12

2 8 9 7 6

7 11 3 9 6

...

page size: 5x8 bytes

DASlab @ Harvard SEAS

when does it make sense to have an oracle
how can we minimize the cost

e.g., **query** x<5

5 10 6 4 12 | 2 8 9 7 6 | 7 11 3 9 6 | ...

algorithm system design = not just computation

CPU DATA MOVEMENT (ENERGY)

MEMORY REQUIREMENT

ROBUSTNESS SPACE REQUIREMENT

SQL, NoSQL, Graph, Neural Nets, Statistics, Vision

CPU DATA MOVEMENT (ENERGY)

MEMORY REQUIREMENT

ROBUSTNESS SPACE REQUIREMENT

SQL, NoSQL, Graph, Neural Nets, Statistics, Vision

CPU DATA MOVEMENT (ENERGY)

MEMORY REQUIREMENT

ROBUSTNESS SPACE REQUIREMENT

TIME —— CLOUD COSTS

Get familiar with the very basics of traditional database architectures:
Architecture of a Database System. By J. Hellerstein, M. Stonebraker and J. Hamilton. Foundations and Trends in Databases, 2007

Get familiar with very basics of modern database architectures:
The Design and Implementation of Modern Column-store Database Systems. By D. Abadi, P. Boncz, S. Harizopoulos, S. Idreos, S. Madden. Foundations and Trends in Databases, 2013

Get familiar with the very basics of modern large scale systems:
Massively Parallel Databases and MapReduce Systems. By Shivnath Babu and Herodotos Herodotou. Foundations and Trends in Databases, 2013

Check out: syllabus, preparation readings, project 0, systems project, online sections

http://daslab.seas.harvard.edu/classes/cs265/

DASlab @ Harvard SEAS

CS 265

Stratos Idreos

BIG DATA SYSTEMS

NoSQL | Neural Networks | SQL | Graph | Data Science