Databases and Big Data

Big Data

Where?

About What?

My Facebook Was Breached by Cambridge Analytica. Was Yours?

How to find out if you are one of the 87 million victims

ROBINSON MEYER APR 10, 2018

Life-size cutouts of Facebook CEO Mark Zuckerberg are displayed by a progressive advocacy group on the lawn of the U.S. Capitol on Tuesday. (CAROLYN KASTER / REUTERS)

MORE STORIES

Mark Zuckerberg Says He's Not Resigning
ROBINSON MEYER

The Cambridge Analytica Scandal, in Three Paragraphs
ROBINSON MEYER

My Cow Game Extracted Your Facebook Data
IAN BOGOST

Sudan and the Instagram Tragedy Hustle
TAYLOR LORENZ

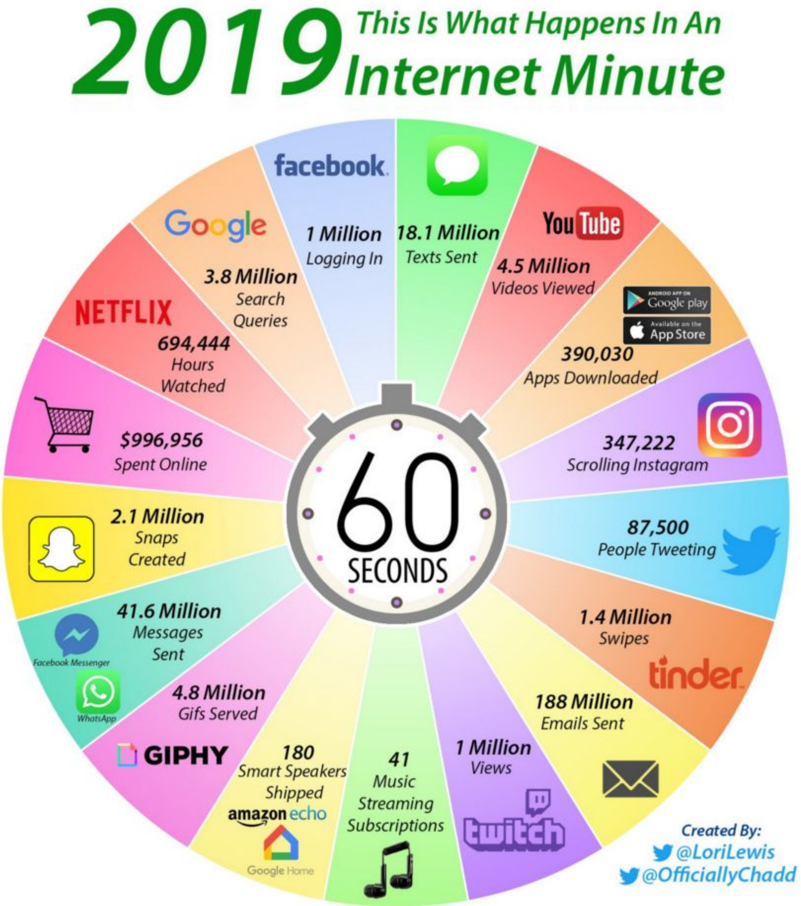
Big Data

- How BIG is big?

kilo (k)	1,000 (3 zeros)
Mega (M)	1,000,000 (6 zeros)
Giga (G)	1,000,000,000 (9 zeros)
Tera (T)	1,000,000,000,000 (12 zeros)
Peta (P)	1,000,000,000,000,000 (15 zeros)
Exa (E)	1,000,000,000,000,000,000 (18 zeros)
Zetta (Z)	1,000,000,000,000,000,000,000 (21 zeros)
Yotta (Y)	1,000,000,000,000,000,000,000,000 (24 zeros)

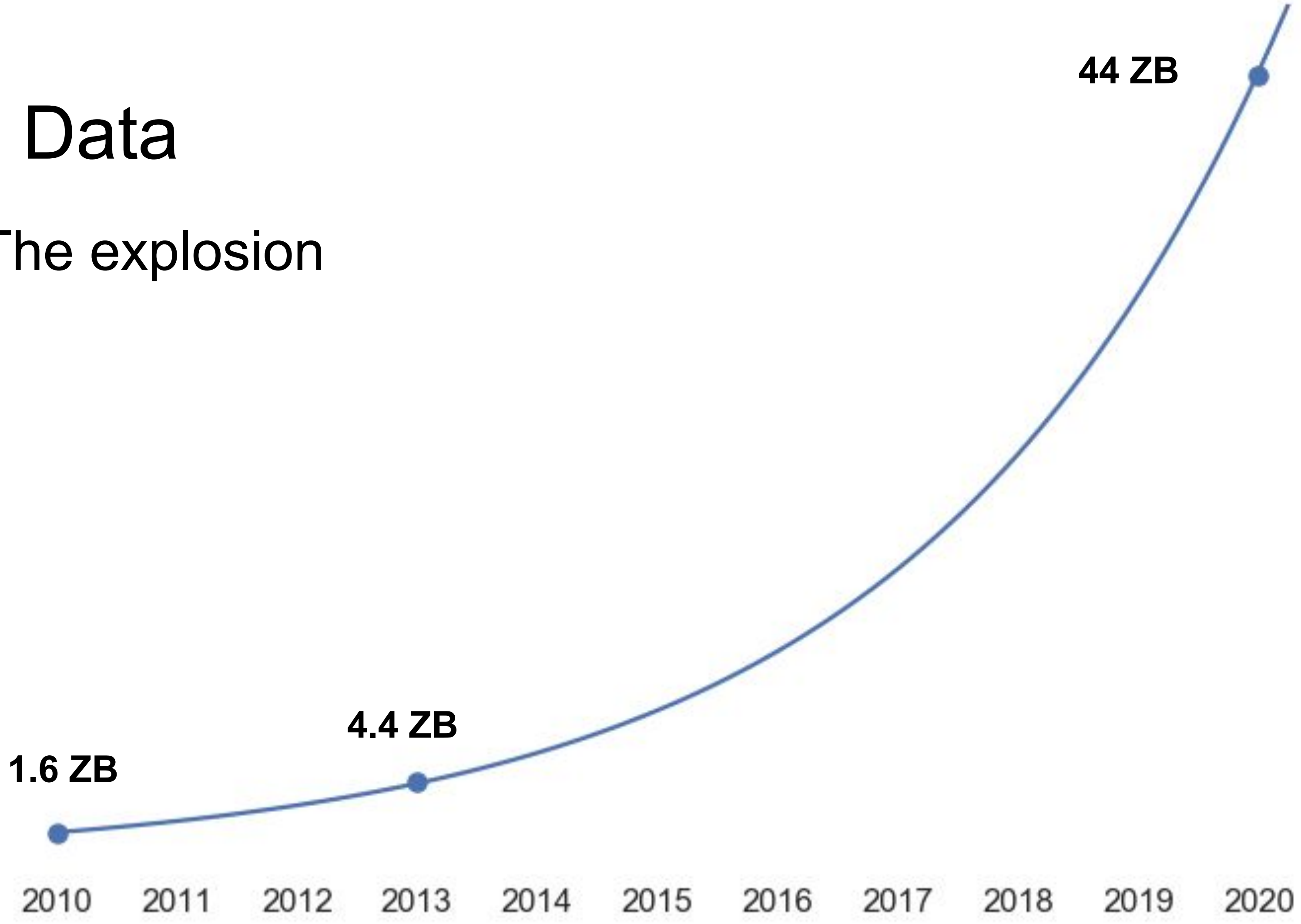
Big Data

- The explosion

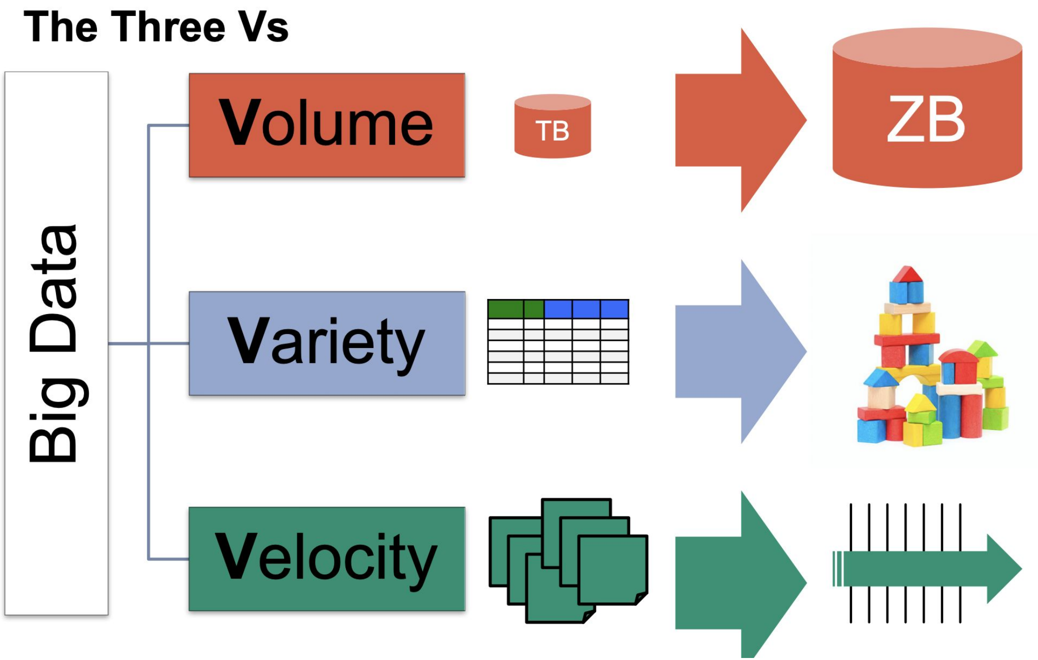


Big Data

- The explosion



Big Data



The Three Vs
Big Data
Volume
TB
ZB
Variety
Velocity

MORE
MORE
MORE

Big Data

- Volume
  - More and more sources
  - Because we CAN!
  - Moore's law on our side

Even better software

Better infrastructure

Bigger, cheaper hardware

Web

Sensor

Science

Big Data

- Volume
    - Because we CAN!
    - Moore's law on our side

Even better software

Better infrastructure

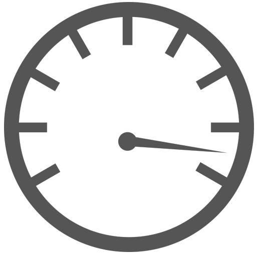
Bigger, cheaper hardware

Big Data

- Volume
    - Because data carries **value**
        - (Most of the time) hard to extract
        - Many people use this to justify keeping more data than needed

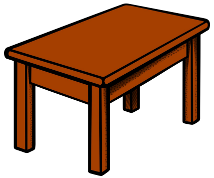
Big Data

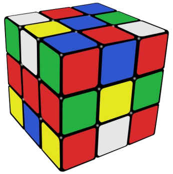
- Velocity
  - Data generated very fast
    - Not manually
    - E.g: 65B whatsapp messages / day
  - You have to handle it in (near) real-time
    - Ingest them all
    - Process them all

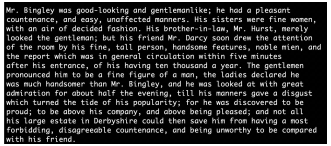


Big Data

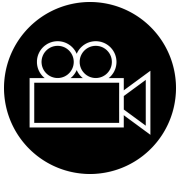
- Variety:
  - All the data models we have learned
  - And more

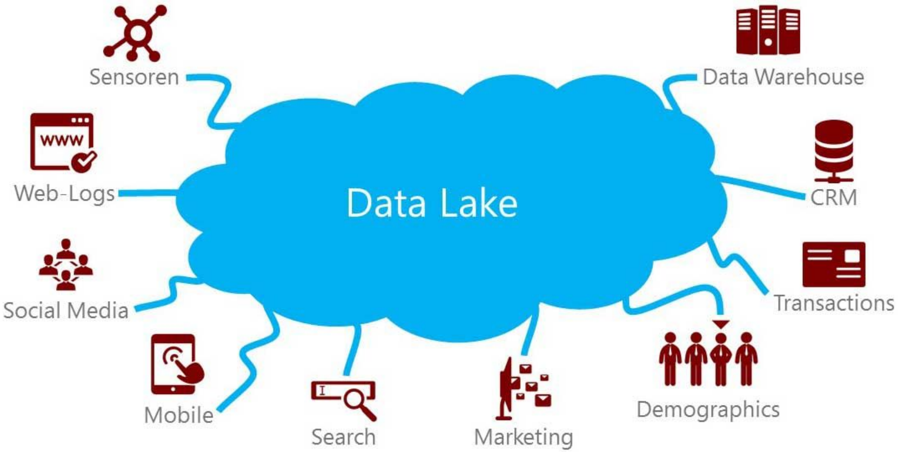






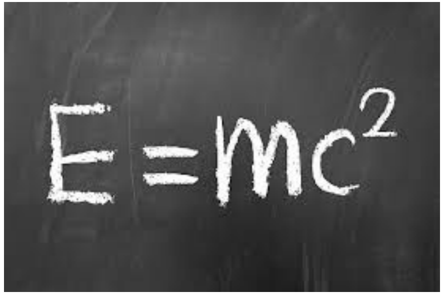
Mr. Bingley was good-looking and gentlemanlike; he had a pleasant countenance, and easy, unaffected manners. His sisters were fine women, with an air of decided fashion. His brother-in-law, Mr. Hurst, merely looked the gentleman; but his friend Mr. Darcy soon drew the attention of the room by his fine, tall person, handsome features, noble mien, and the report which was in general circulation within five minutes after his entrance, of his having ten thousand a year. The gentlemen pronounced him to be a fine figure of a man, the ladies declared he was much handsomer than Mr. Bingley, and he was looked at with great admiration for about half the evening, till his manners gave a disgust which turned the tide of his popularity; for he was discovered to be proud; to be above his company, and above being pleased; and not all his large estate in Derbyshire could then save him from having a most forbidding, disagreeable countenance, and being unworthy to be compared with his friend.





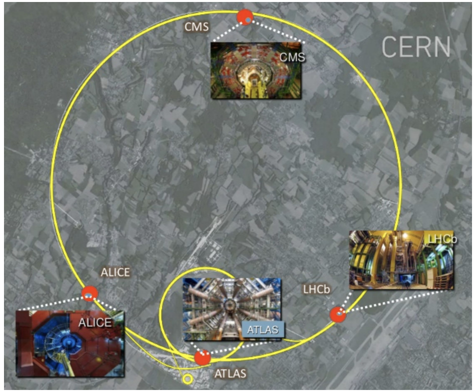
Why Big Data

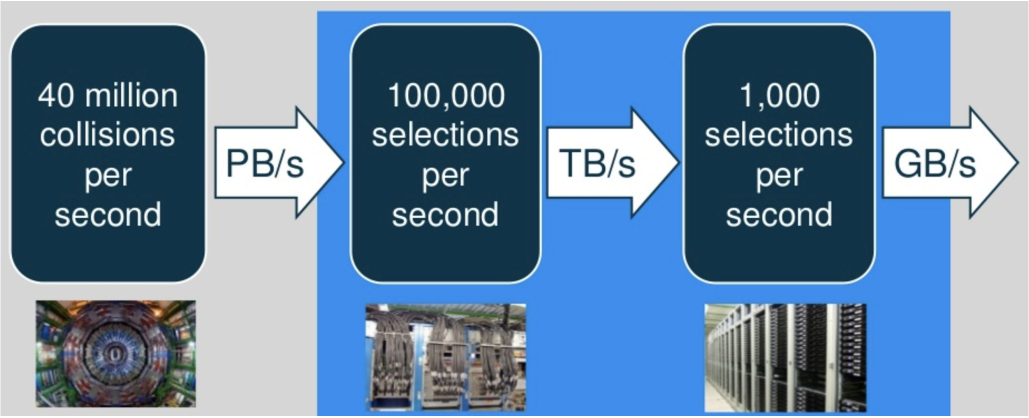
- Because we HAVE TO
- 4th scientific paradigm
  - 1st: observation
  - 2nd: theoretical (make hypothesis)
  - 3rd: computational
  - **4th: data-driven**

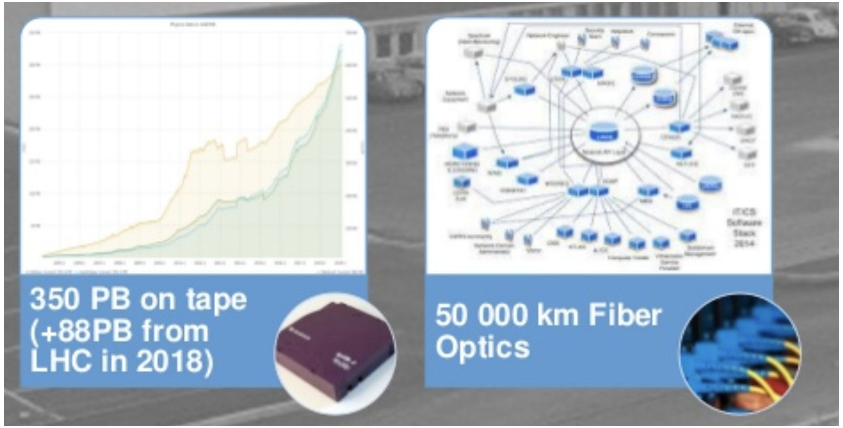


Big Data

- Data-Driven Science

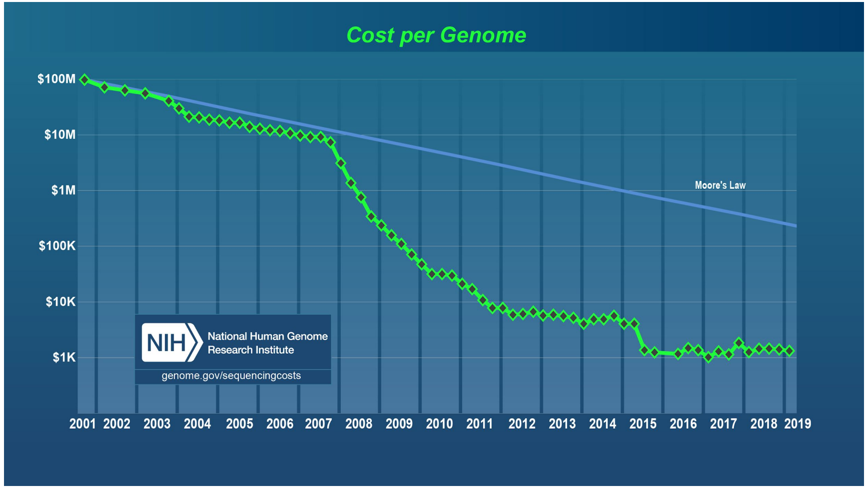




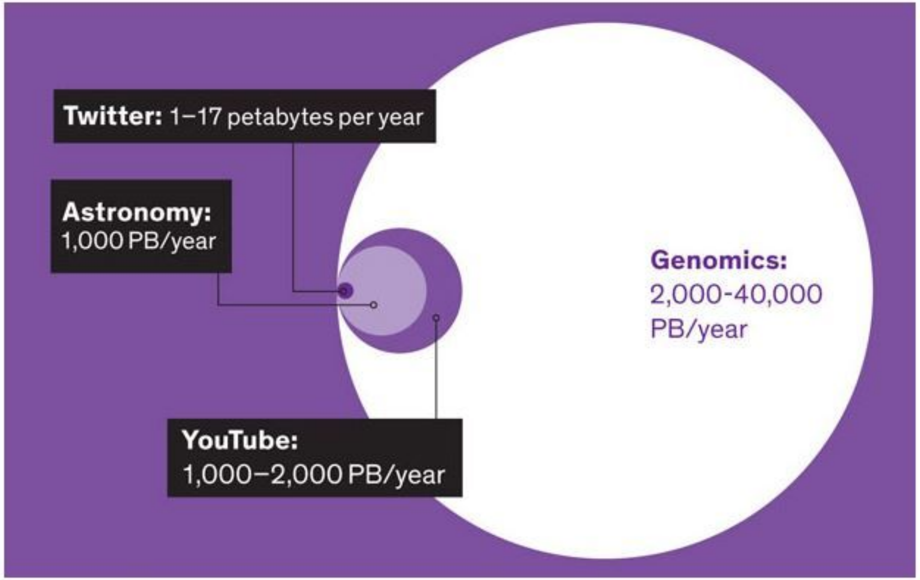


Big Data

- Data-driven science



Projected annual storage in 2025



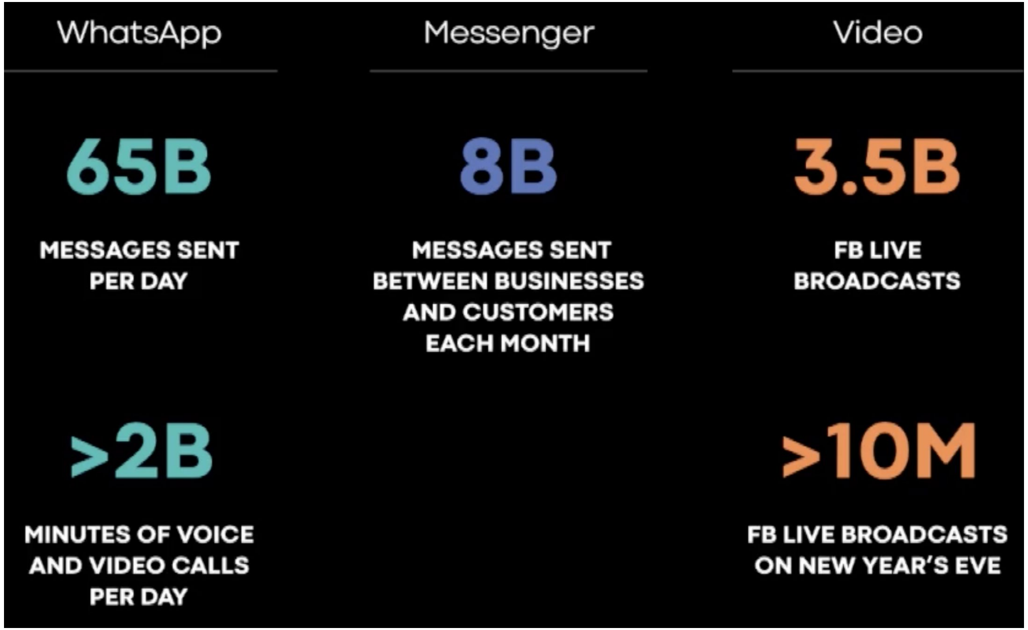
Source: "Big Data: Astronomical or Genomical?" PLoS Biology, 7 July 2015.

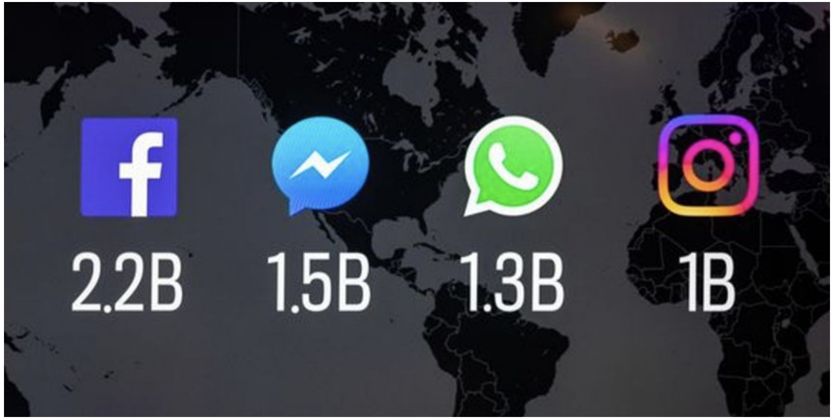
Big Data

- Finance

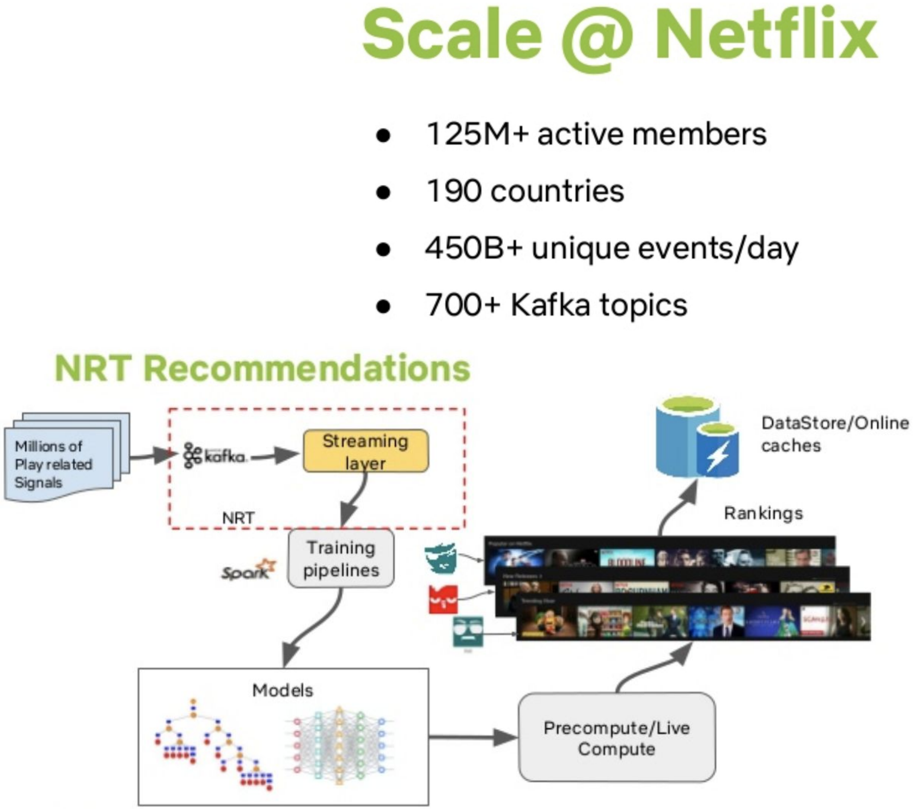
Big Data

- Our lives





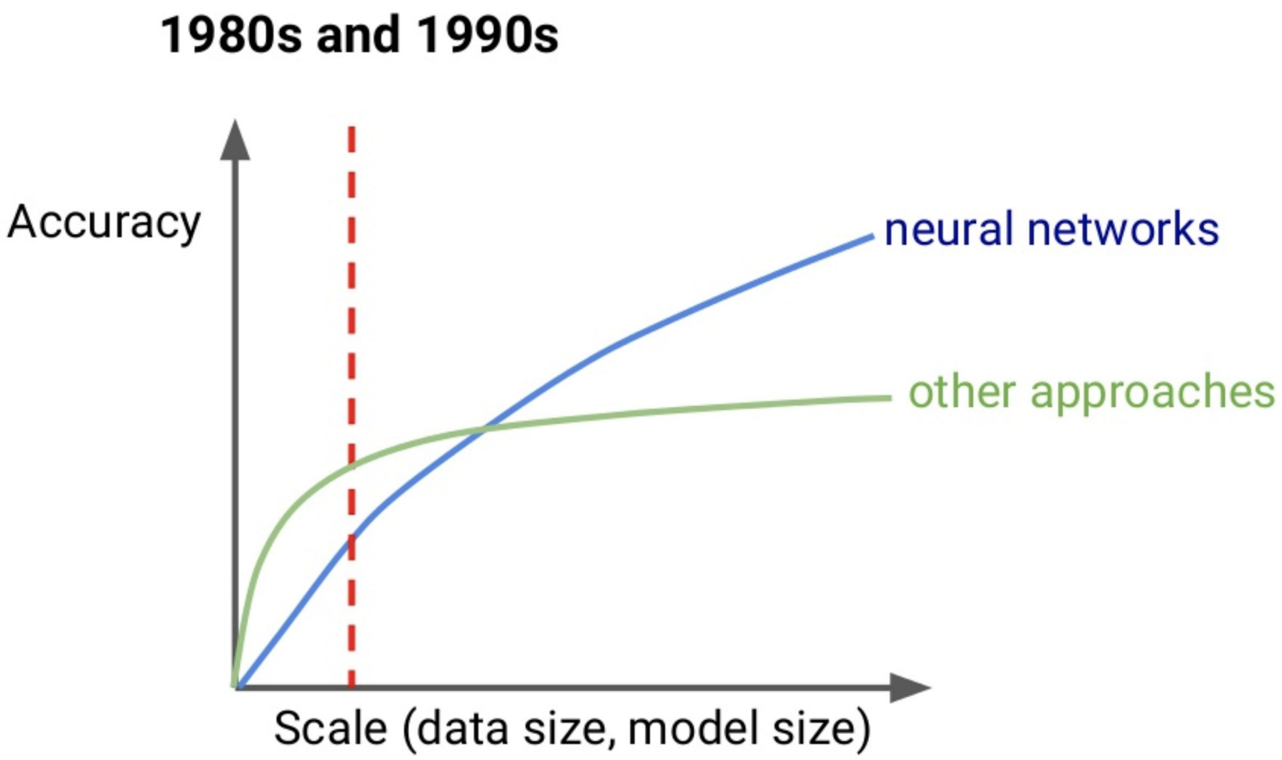
Big Data - Impact



Big Data

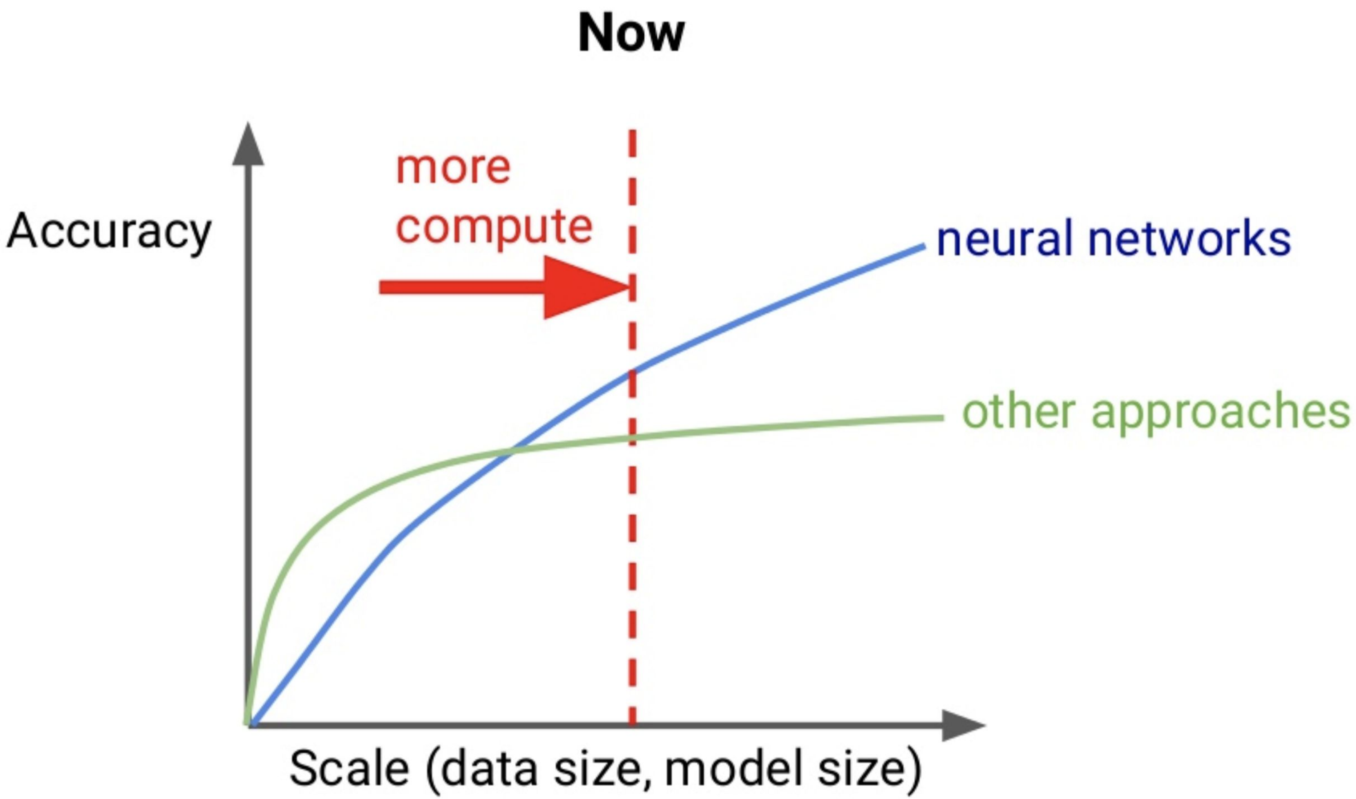
- Machine learning

Big Data



1980s and 1990s
Accuracy
neural networks
other approaches
Scale (data size, model size)

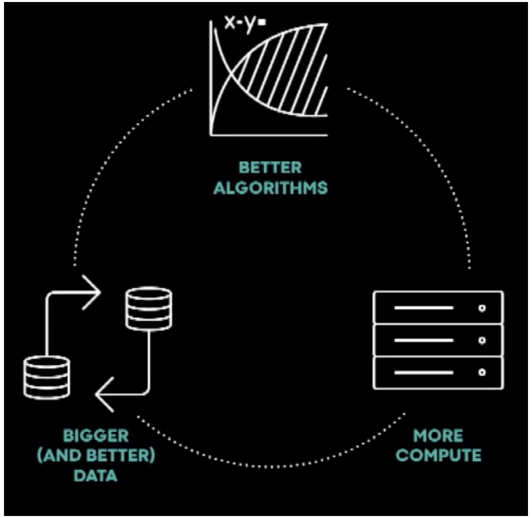
Big Data

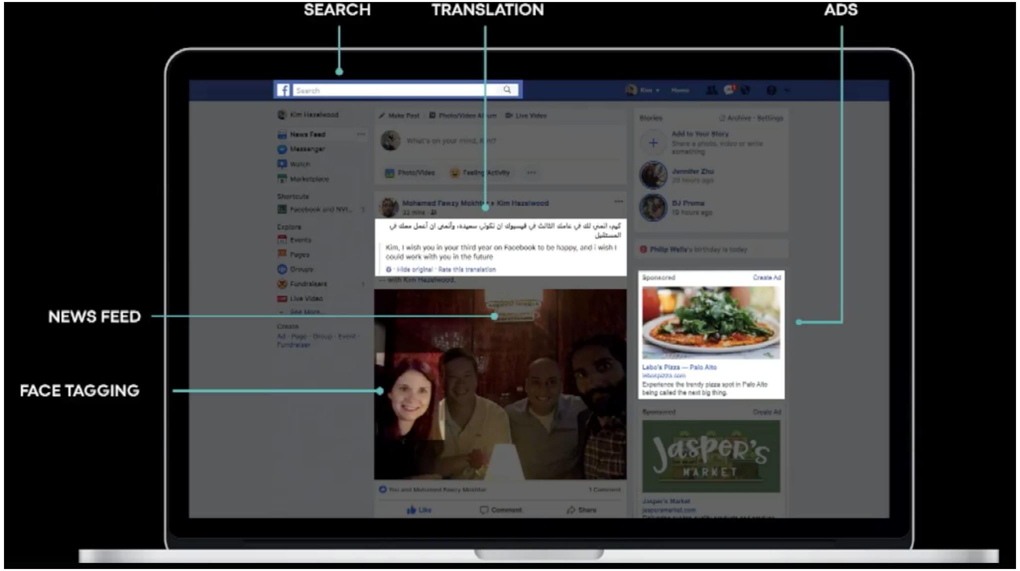


Now
more
compute
Accuracy
neural networks
other approaches
Scale (data size, model size)

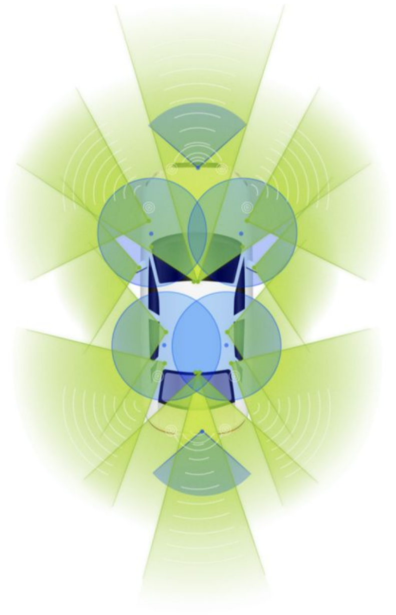
Big Data - Impact

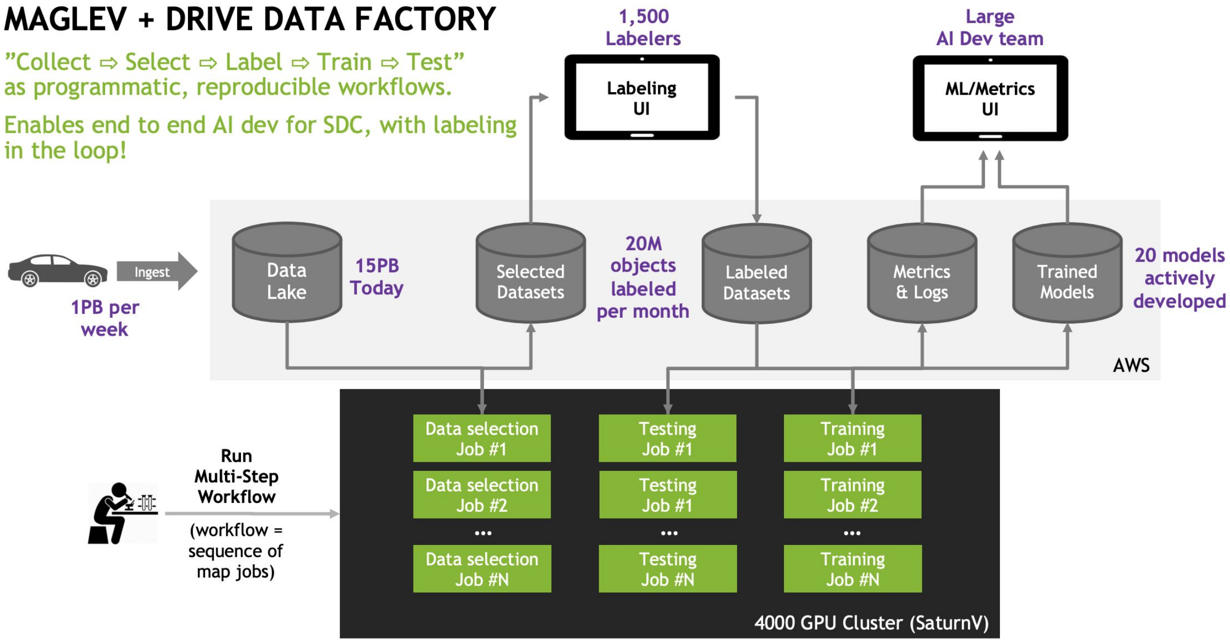
- Virtuous cycle
- ML becomes ubiquitous





Big Data - Self Driving Car



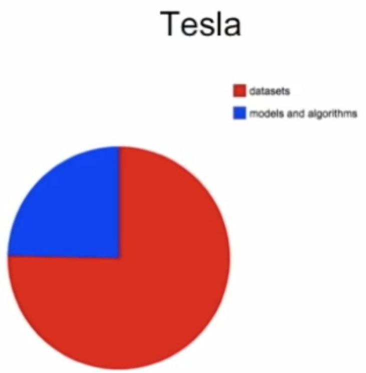


NVIDIA

Big Data

Largest deployment of robots in the world (0.25M)

Make them autonomous.



Big Data - Impact

- Not all good

Robust De-anonymization of Large Sparse Datasets

Arvind Narayanan and Vitaly Shmatikov
The University of Texas at Austin

Systems' Perspective

- Remember 3 V's:
    - Volume, Velocity, Variety

Does **not** work at
BIG DATA scale

Volume

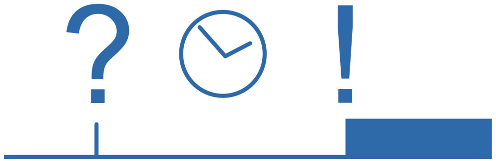
Velocity

Variety

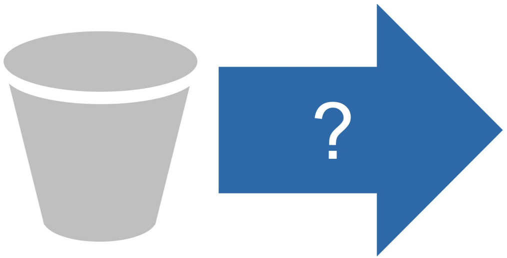
Big Data Systems

- How to judge a big data system
    - Capacity
    - Throughput
    - Latency

How much we can store?



How fast we can process a record/request?

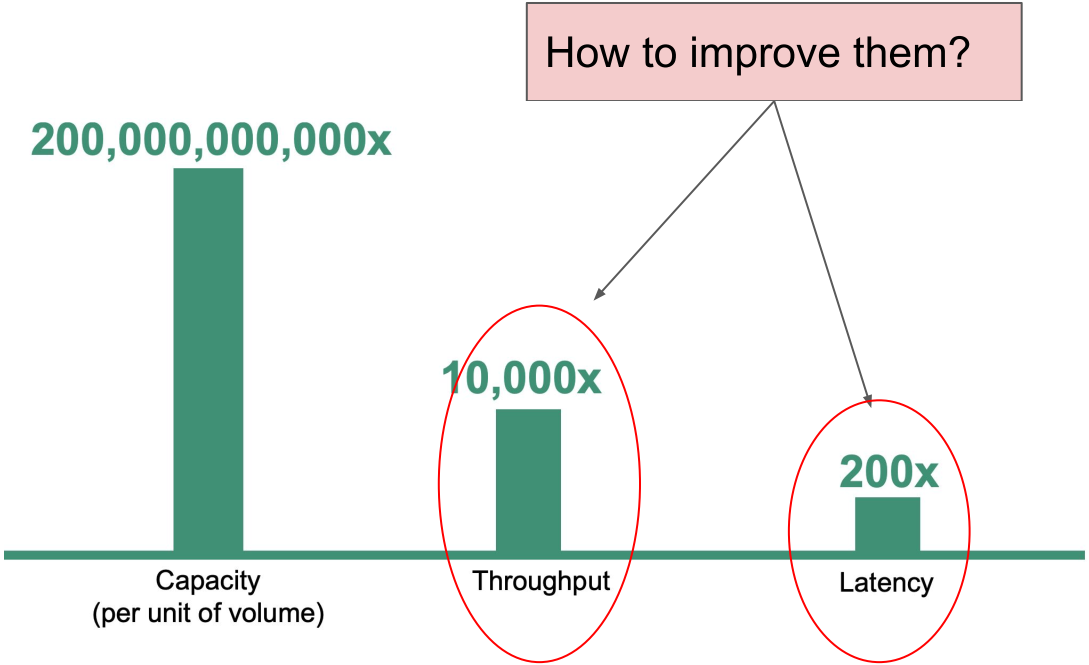


How much we can process per second?

Big Data Systems

- Moore's law is not equal

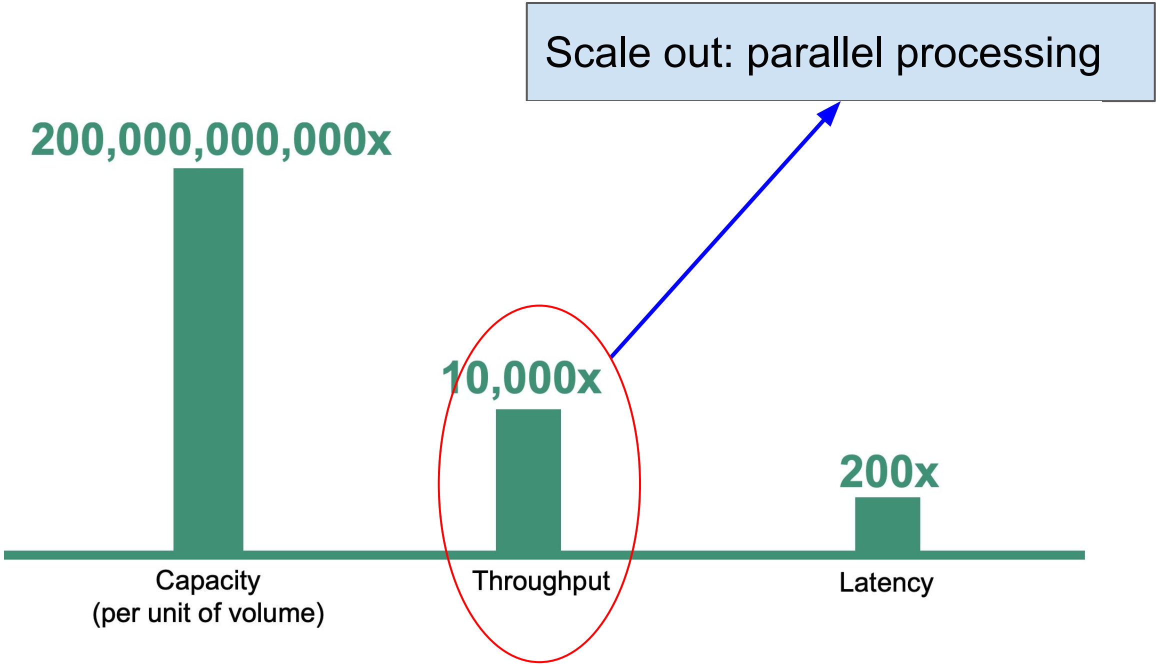
Disk improvement
(log scale)
1960 - 2018



Big Data Systems

- Moore's law is not equal

Disk improvement
(log scale)
1960 - 2018



Process in parallel since your throughput is 1000 record/s
Worst latency you can have is 10ms too
(avg latency gone down)
as compared to when you process in the queue one by one
(avg latency is very high)

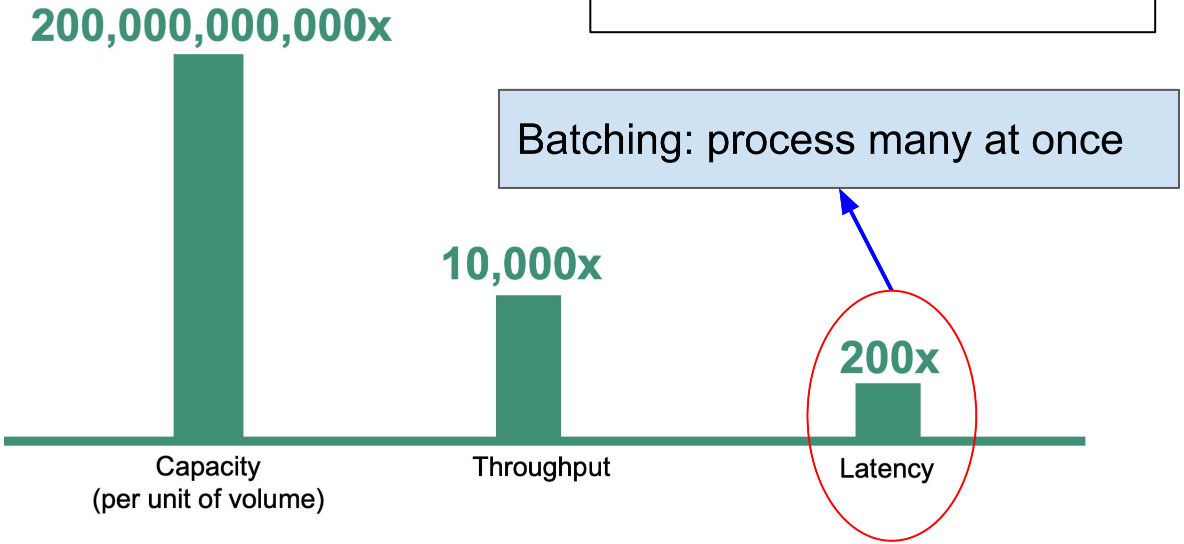
Big Data Systems

- Moore's law is not equal

Your system:
  Latency: 10ms per record
  Throughput: 1000 records/second

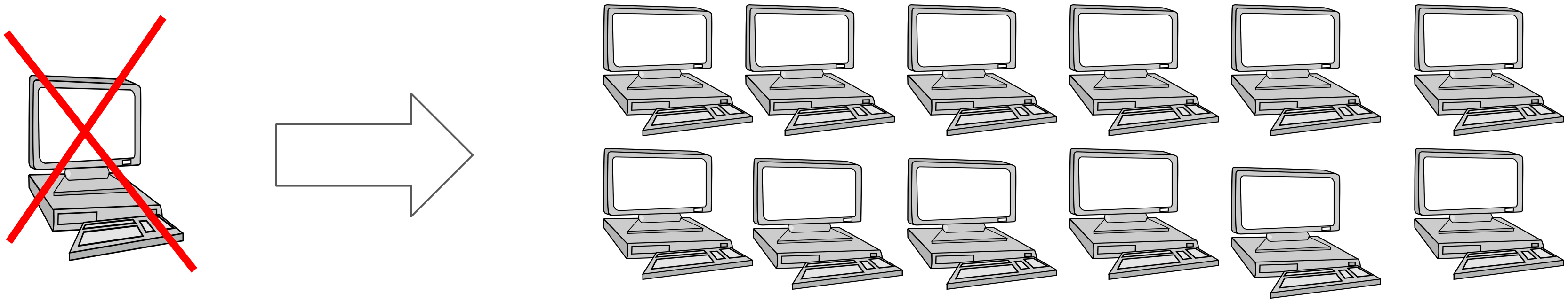
1000 records arrive at the same time
What do you do?

Disk improvement
(log scale)
1960 - 2018



Big Data Challenges

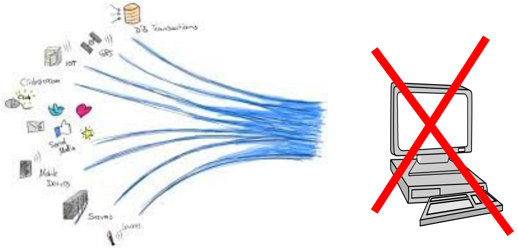
- Challenge 1: **volume**

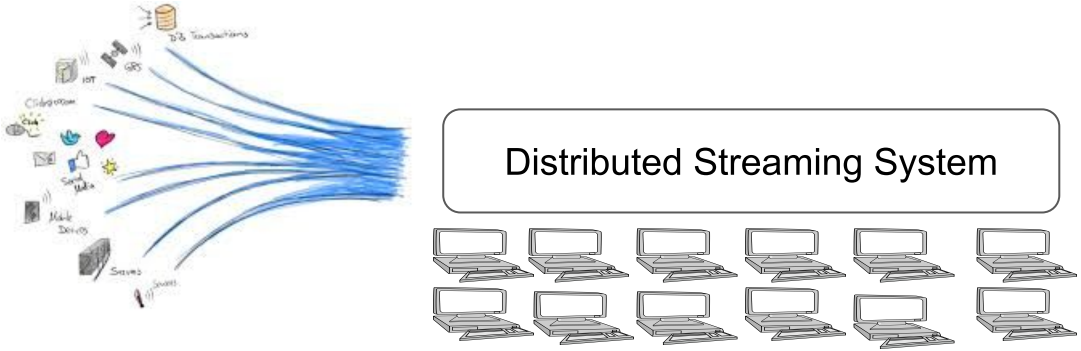


Approach: Scale OUT

Big Data Challenges

- Challenge 2: **velocity**

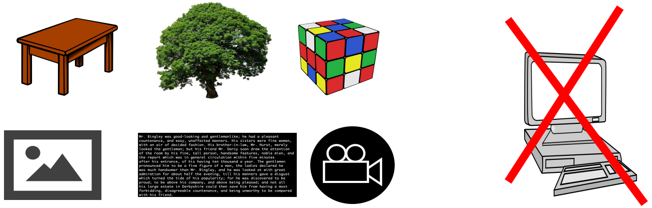




Approach: Scale out + Streaming systems

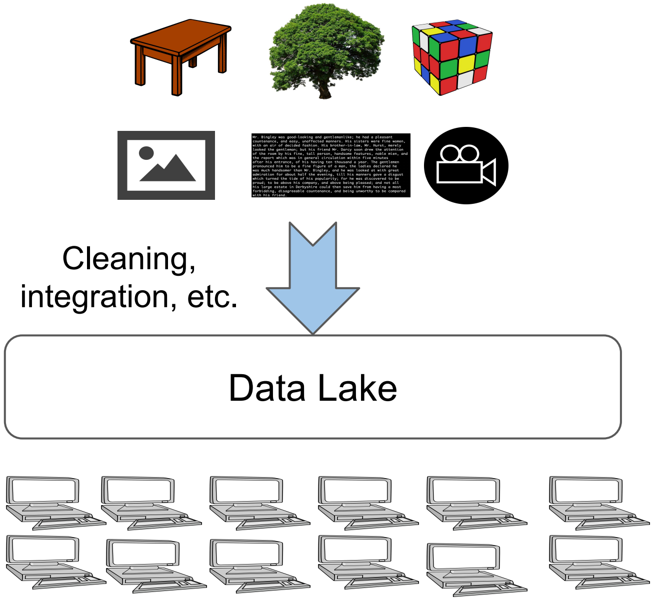
Big Data Challenge

- Challenge 3: **variety**



Approach: collect now, figure out later

Specialised system: *may work and better perf but not scalable (not sustainable) cause new data type always come up*



Approach: specialized system

What is Big Data

- Not only about the data
- An ecosystem (suite of technologies) for:
  - Storing
  - Managing
  - Analyzing

Data that is too big to fit on a single machine

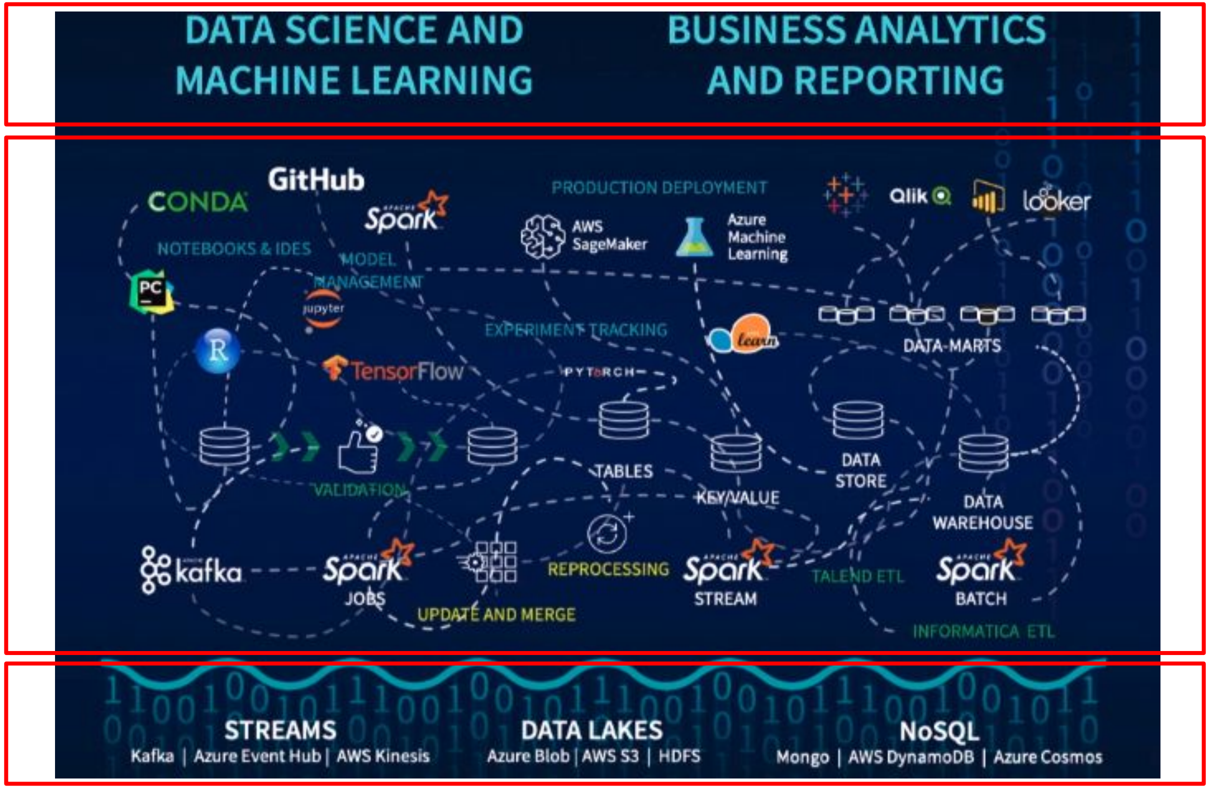
At scale

Performance
Fault tolerance
Scalability

Applications

Systems

Data



BIG DATA & AI LANDSCAPE 2018

V1 – Last updated 6/19/2018

© Matt Turck (@mattturck), Demilade Obayomi (@demi_obayomi), & FirstMark (@firstmarkcap) mattturck.com/bigdata2018

FIRSTMARK
EARLY STAGE VENTURE CAPITAL

Big Data Storage

- Dynamo
  - 2007
  - Behind Amazon's S3
  - One of the first big data system

amazon

Key-value

<customer, profile>
...

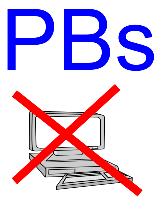
Dynamo: Amazon's Highly Available Key-value Store

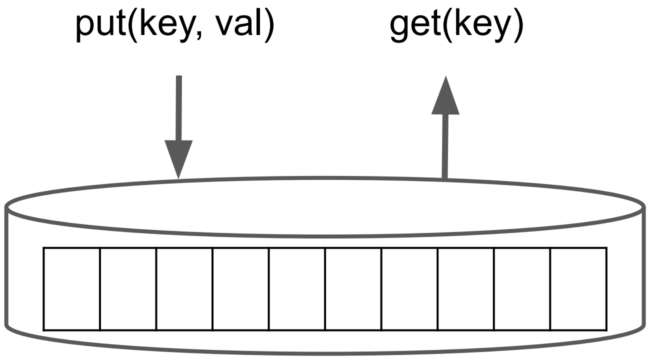
Giuseppe DeCandia, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati, Avinash Lakshman, Alex Pilchin, Swaminathan Sivasubramanian, Peter Vosshall and Werner Vogels

Amazon.com

Dynamo

- Problems it tries to solve:
    - Store many, many key-value records
        - Doesn't fit on a single machine
    - Efficient search based on key

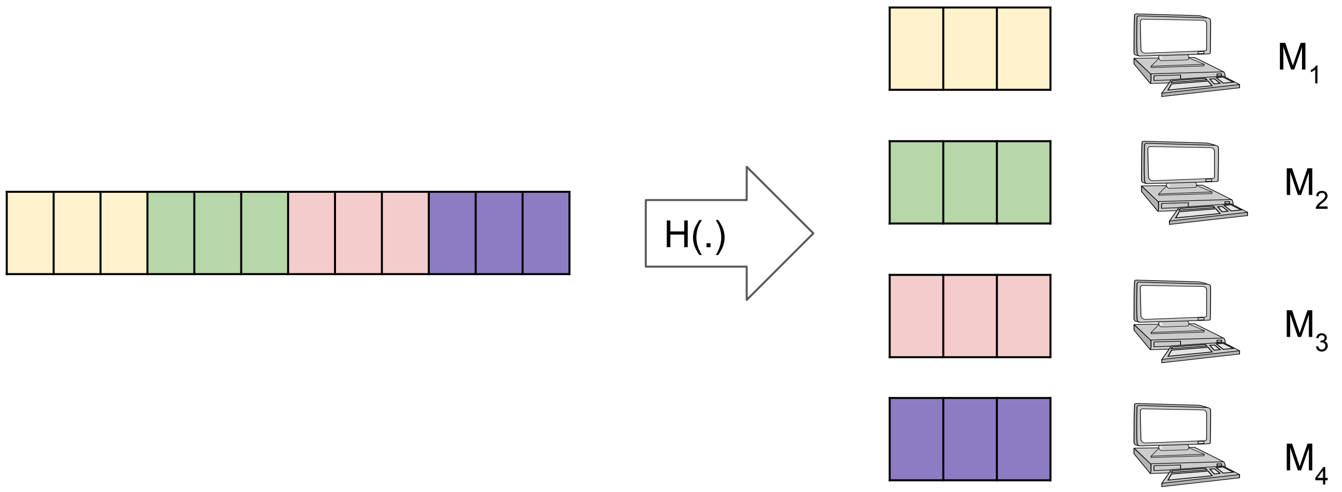




One
Giant
Hash
Table

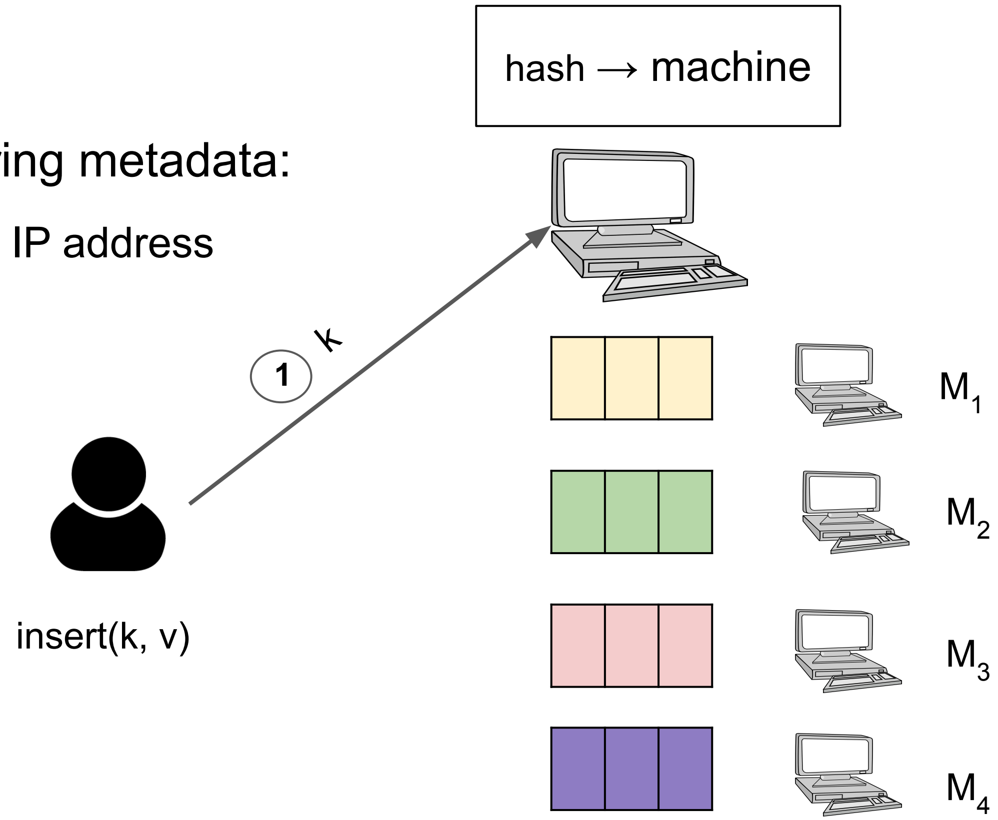
Dynamo

- Strawman
  - Chop up the giant hash table
  - Store pieces on different machines



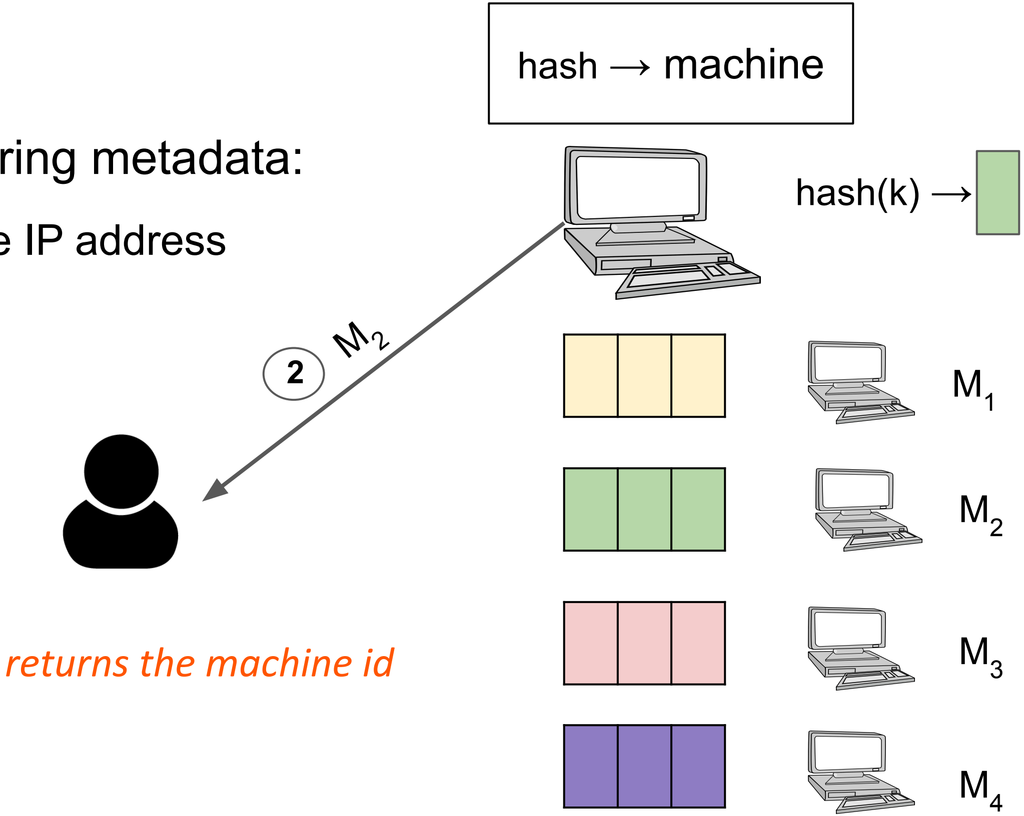
Dynamo

- Strawman
  - One master machine storing metadata:
    - Hash value → machine IP address



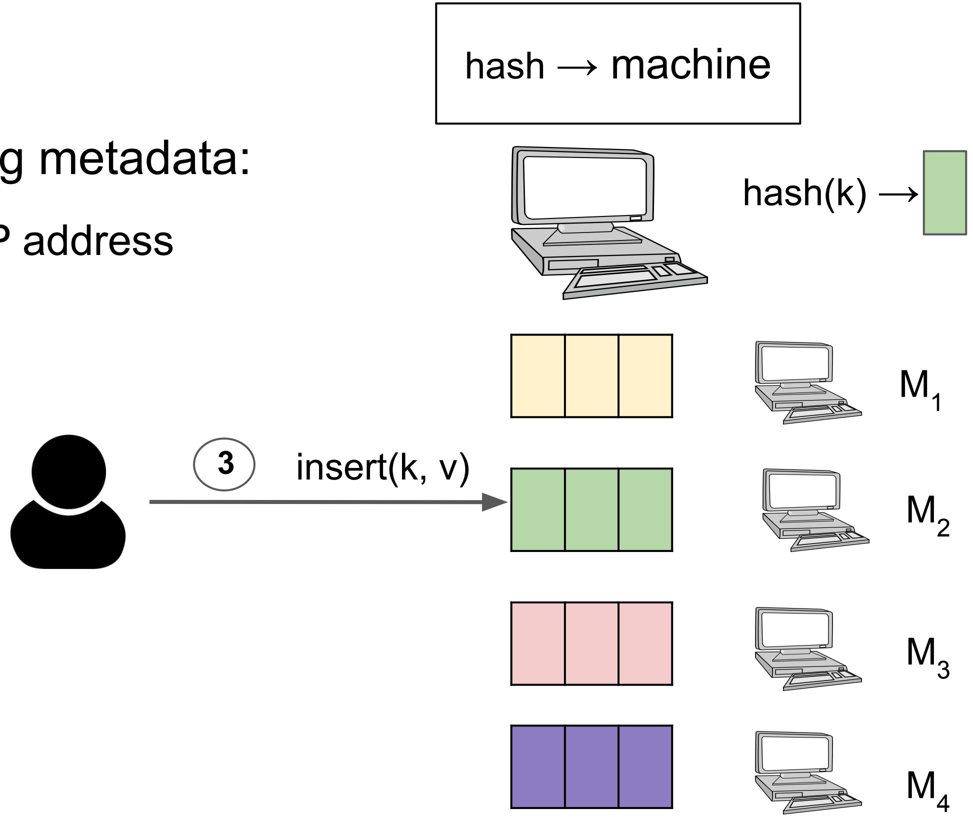
Dynamo

- Strawman
  - One master machine storing metadata:
    - Hash value → machine IP address



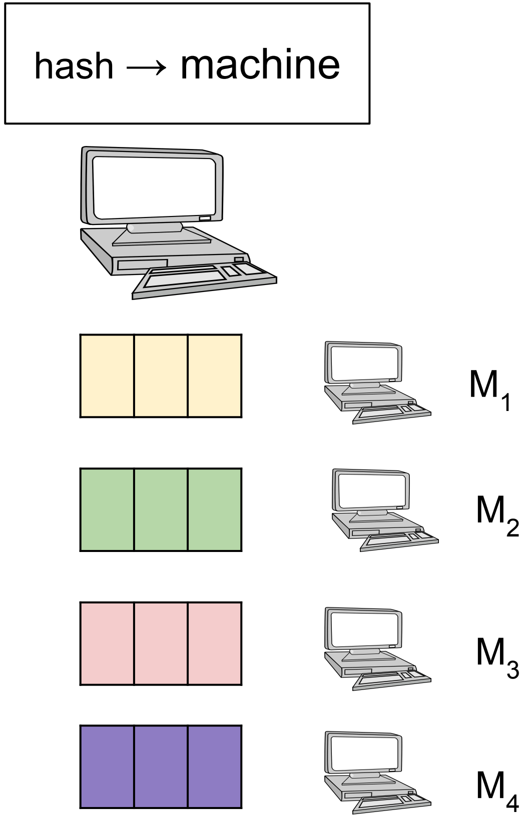
Dynamo

- Strawman
  - One master machine storing metadata:
    - Hash value → machine IP address



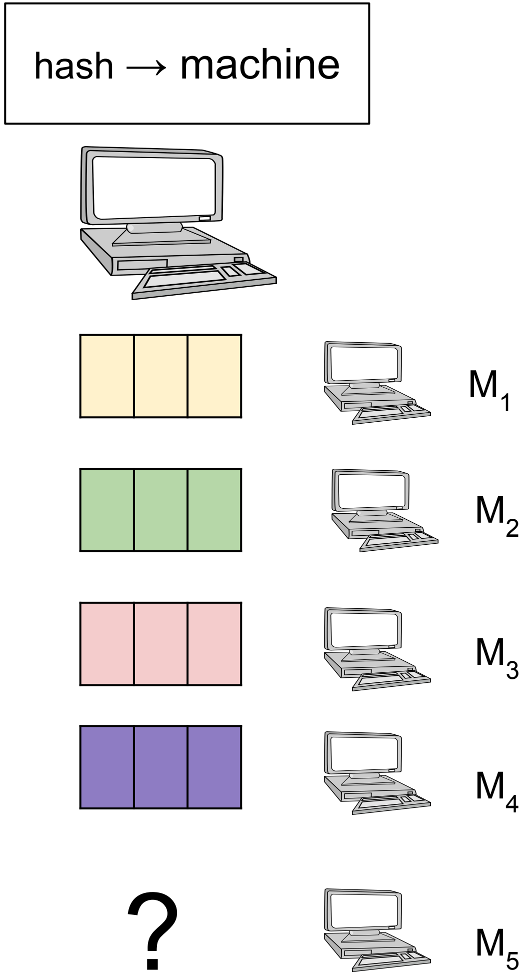
Dynamo

- Strawman
  - Problem:
    - What happen if I want to add new machine?
    - (or if one machine fails)?



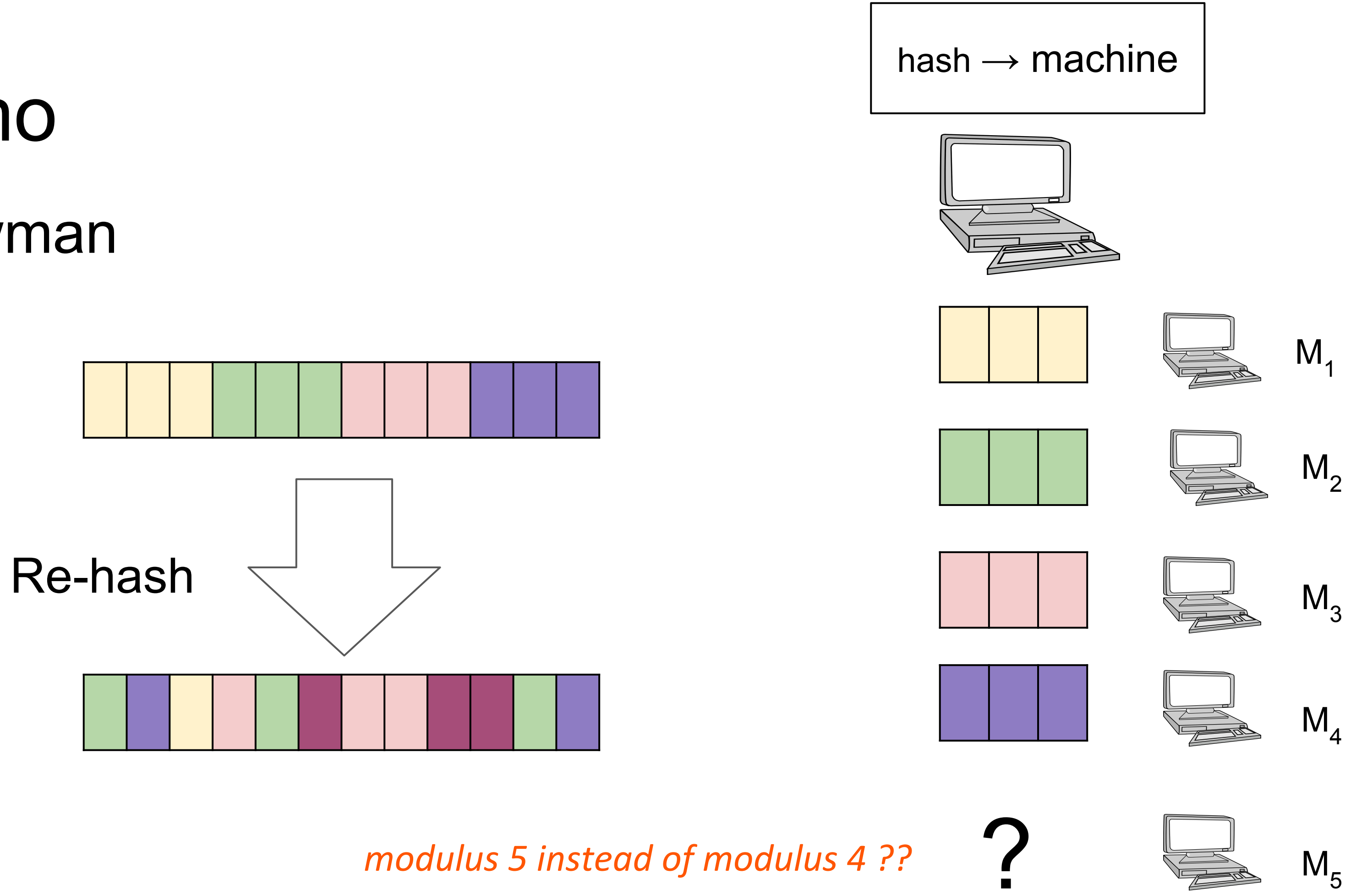
Dynamo

- Strawman



Dynamo

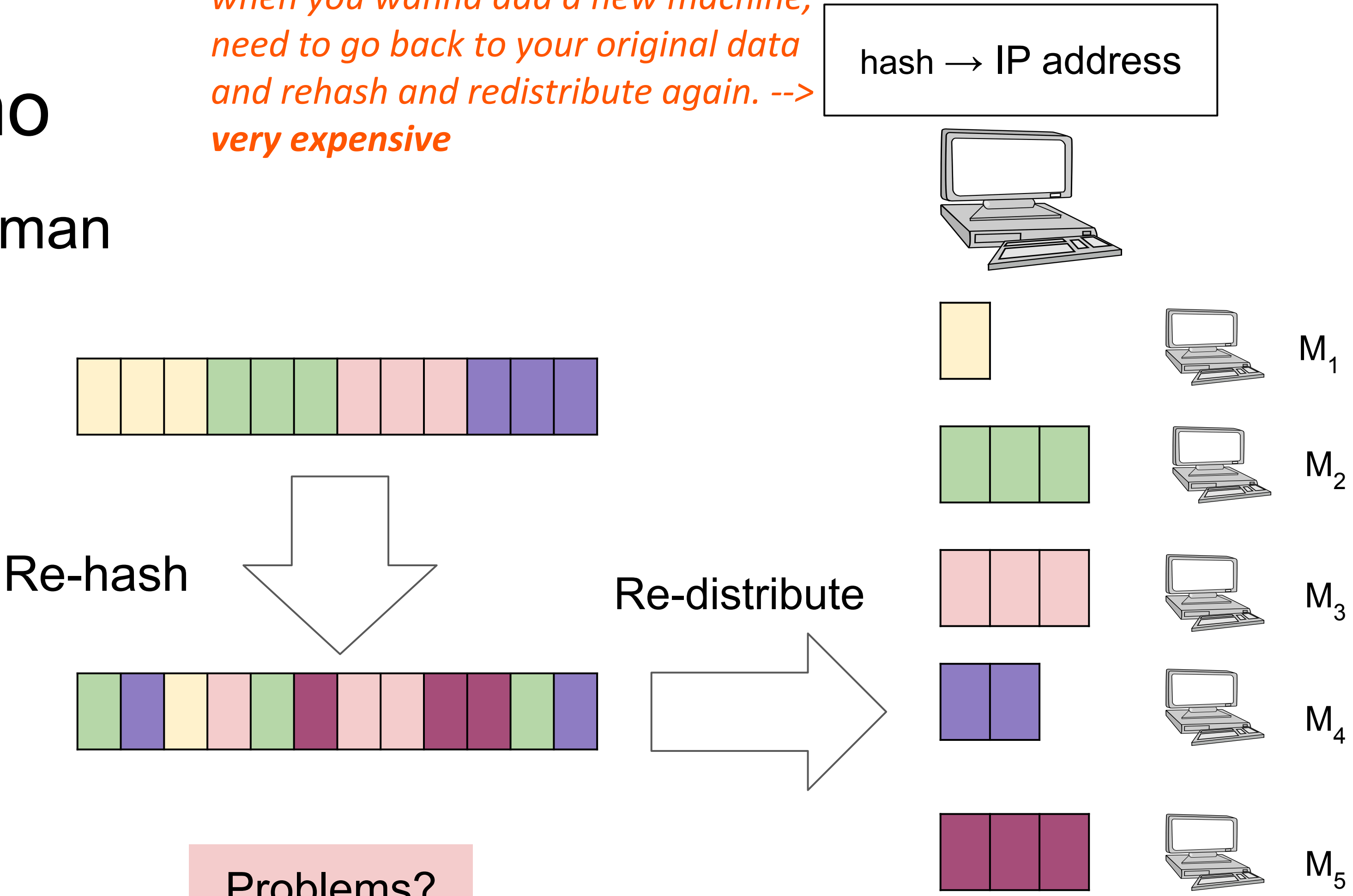
- Strawman



Dynamo

*when you wanna add a new machine, need to go back to your original data and rehash and redistribute again. --> **very expensive***

- Strawman



Problems?

Dynamo

- Consistent Hashing
  - Machines and keys get their IDs from the same space:

    $[0,..,2^m)$

m = 6 → ID space: 0..63

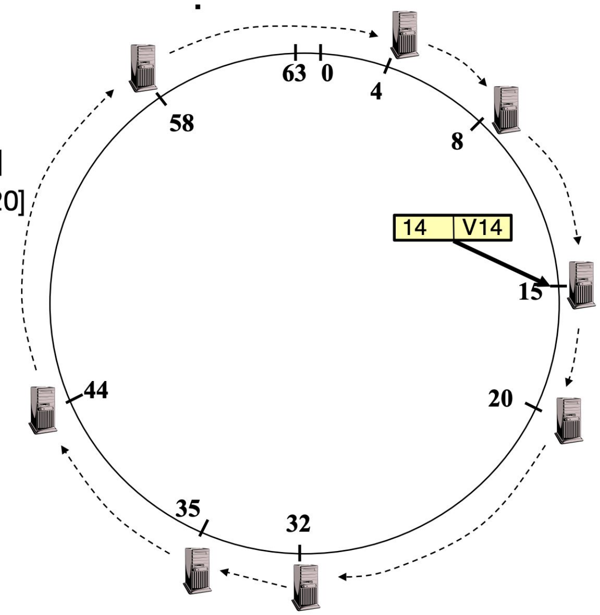
Node 8 maps keys [5,8]

Node 15 maps keys [9,15]

Node 20 maps keys [16, 20]

…

Node 4 maps keys [59, 4]

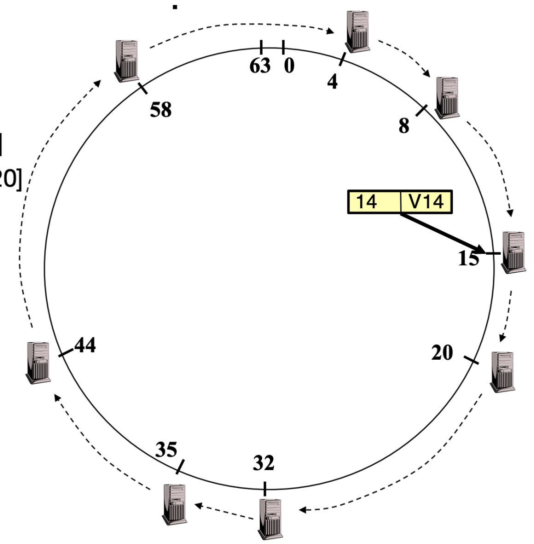


Very important idea in modern computing systems

Dynamo

- Consistent Hashing:
  - IDs is a circle
  - Node stores only keys in the range between itself and the previous node (clockwise)

m = 6 → ID space: 0..63
Node 8 maps keys [5,8]
Node 15 maps keys [9,15]
Node 20 maps keys [16, 20]
…
Node 4 maps keys [59, 4]



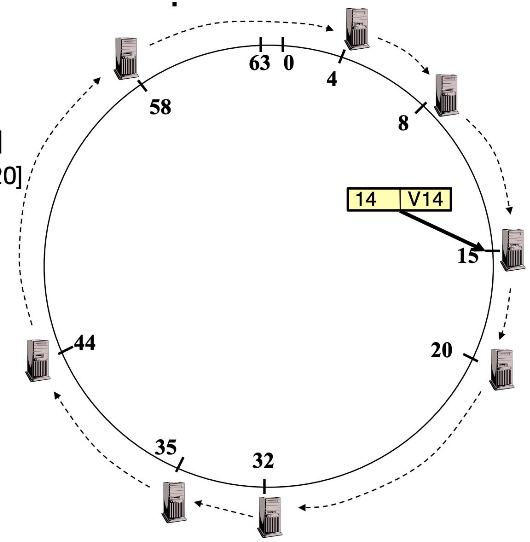
Node 15 stores keys in the range (8, 15]

Node 32 stores keys in the range (32,35]

Dynamo

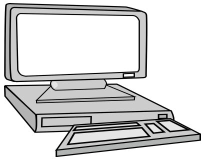
- Consistent Hashing:
    - Example: `insert(30, val)`
        - `val` stored at node 32

m = 6 → ID space: 0..63
Node 8 maps keys [5,8]
Node 15 maps keys [9,15]
Node 20 maps keys [16, 20]
…
Node 4 maps keys [59, 4]



Node ID → IP address

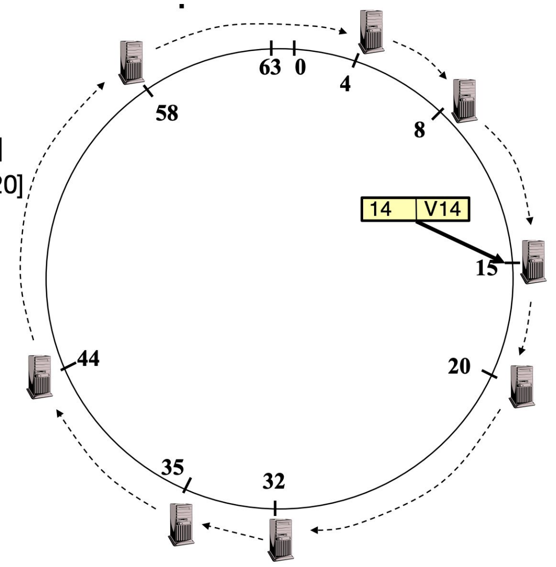
Master server



Dynamo

- Consistent Hashing:
  - Why better?
  - What happen when a new node joins?
    - Get a new node ID
    - Then….?

m = 6 → ID space: 0..63
Node 8 maps keys [5,8]
Node 15 maps keys [9,15]
Node 20 maps keys [16, 20]
…
Node 4 maps keys [59, 4]

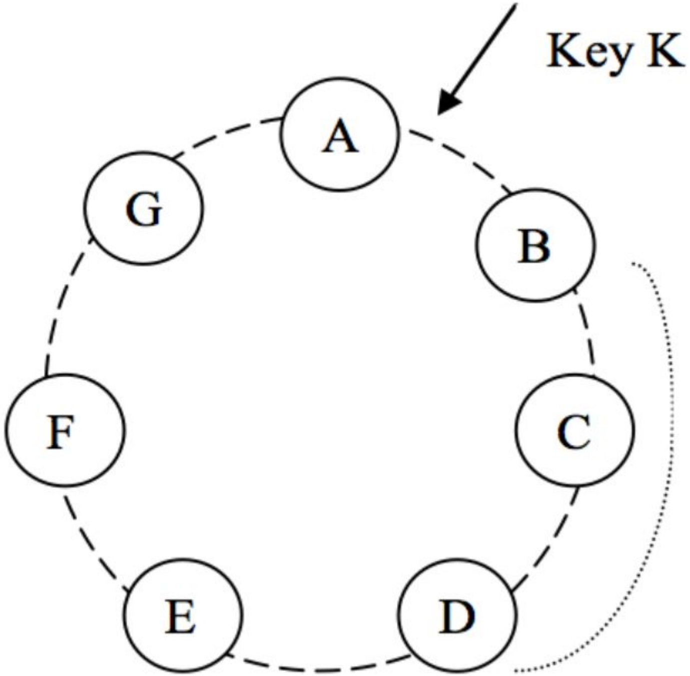


insert node 40
only need to redistribute key range in node 44
if node 44 fails, the only node that is affected is node after it.

if bottle neck at node b, you replicate and get the same data into node c and d.

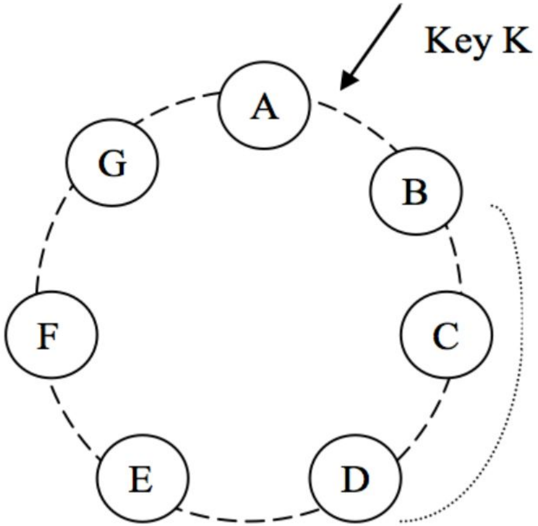
Dynamo

- Replication:
  - Usually 3 successive nodes in the ring
    - Not physically close to each other
  - E.g. node B, C, D stores K
- Write operation
  - Sent to any of these replicas
- Read
  - From any of these replicas



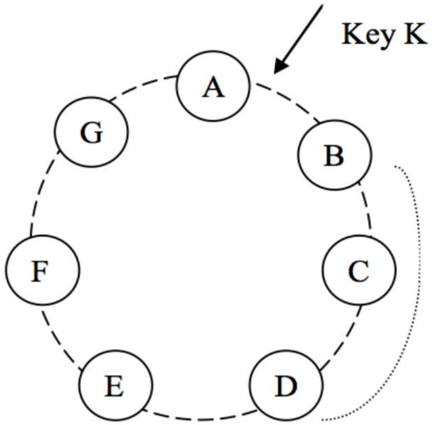
Dynamo

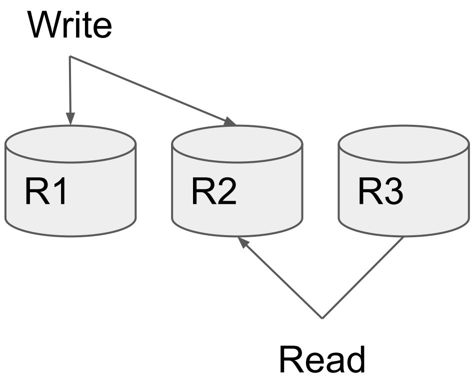
- Problem: eventual consistency
    - Example:
        - Write to B
        - B starts replicating to C and D
        - Read from D, when previous step hasn't finished
    - Eventual consistency: no guarantee that read returns the latest write
    - What happen when 2 users read from 2 different replicas?



Dynamo

- Problem: eventual consistency
  - Amazon say it's OK!
    - User can add to items to cart, only later find out that item no longer available
  - Sacrificing correctness for performance
- A middle ground: sloppy quorum
  - N replicas
  - Write waits for W replicas to finish (W < N)
  - Read waits for R replicas to finish (R < N)

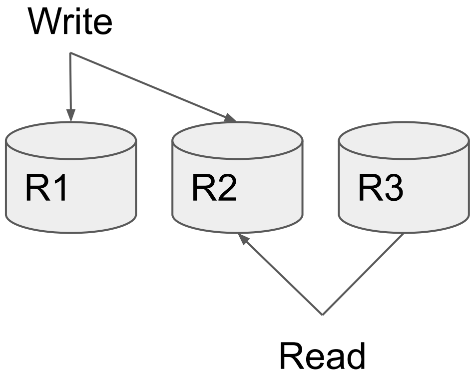




Dynamo

- A middle ground: sloppy quorum
  - N replicas
  - Write waits for W replicas to finish (W < N)
  - Read waits for R replicas to finish (R < N)
- When R + W > N
  - Set of write and read replicas overlap at 1 least 1 server
  - Meaning that at least 1 replica contains the **<u>latest write</u>**

How do we know which one?



Example: N=3
- W = 2
- R = 2

Summary

- Big data is
  - Volume, Velocity, Variety
  - Many specialized systems put together
- Making real impact
  - Necessity, no longer a choice, for large enterprises
- Few examples: key-value store, stream, graph
- Next: Cloud computing
  - Key enabler of Big data