SOUNDCLOUD

FITTING THE MODEL

Oct 3rd, 2012

Sean Treadway

Hej, Sean ★ | Settings | Inbox 1 | Log out

App Gallery | Get help | Try Premium | What's new?

Dashboard | You | Tracks | People | Groups | Upload & Share

Search SoundCloud

Dashboard

Your activity 10+ | Incoming tracks 10+ | Exclusive tracks 1

moltke **uploaded a new track**

QR-Grabsteine in Dänemark (für Breitband, D-Radio Kultur)

moltke 1 day ago

55 ▶ | 1 💬

Share | Save to Favorites | Download | Radio

0.00 / 5.33

Eric ● **uploaded a new track**

melodic done drill at SoundCloud ReBase

Eric ● 1 day ago

Share | Save to Favorites | Download

0.00 / 0.31

Tanith **uploaded a new track**

Sommermix2012- Part 2

Tanith 1 day ago

257 ▶ | 3 💬 | 16 ♥ | 60 ⬇

Share | Save to Favorites | Download | Bass

Mini Update | More

	This Week	Total
Plays	15	608
Comments	-	44
Favoritings	-	11
Downloads	-	66
Profile Views	128	22509

Following 121 | Followers 1028 10+ | Groups 1

SHIGETO ★ ●

48 | 23110 — Following

Johannes Wagener ★ ●

63 | 1365 — Following

Kimbramusic ★ ●

18 | 4916 — Following

Hannes ★ ●

65 | 9751 — Following

Stream | You | Explore | Upload | Info

Your Stream

Dam Mantle ↻ seams — 15 hours

Brothers Fowl (LP preview mix)

+ | ↻ | ♥ | Share | ↓ | Buy

▶ 239 | 2 | ↻ 1 | ♥ 10

DW - Learn German — 8 hours

Bierselig | Sprachbar | Sep 19, 2012

+ | ↻ | ♥ | Share | ↓

2:54 Band ↻ Eric — 18 hours

2:54 - Killer

Play on!

Explore while listening to sounds non-stop with the new Continuous Play feature.

Who to follow — More

lauraherberg
7 | 5942 — Follow

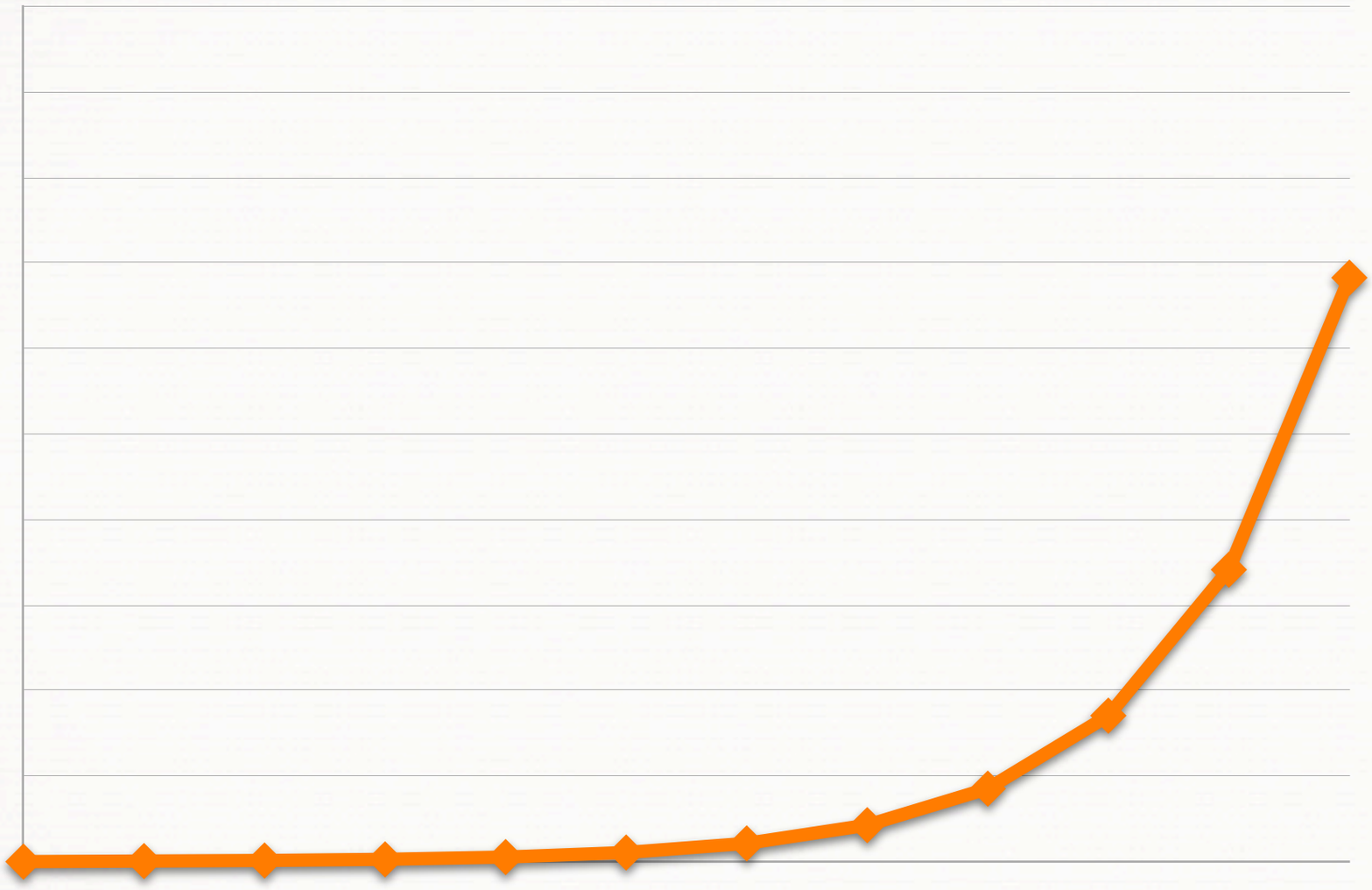
TheGaslightAnthem
57 | 11754 — Follow

WorkingNow
8 | 18309 — Follow

Recent activity — View all

C.A ELITE 2012 followed... — 11 hours
3 | 88 — Follow Back

GROWTH



Stuff we Serve

DATABASE

Engineering is the art of modeling **materials** we do not wholly understand, into shapes we cannot precisely analyse so as to withstand forces we cannot properly assess, in such a way that the public has no reason to suspect the extent of our ignorance.

Dr A. R. Dykes - British Institution of Structural Engineers, 1976.

MYSQL → CASSANDRA

(does it matter?)

PRIMARY / SECONDARY

SPINNING DISKS

Dashboard

moltke **uploaded a new track**

QR-Grabsteine in Dänemark (für Breitband, D-Radio Kultur)

moltke 1 day ago

55 ▶ | 1 💬

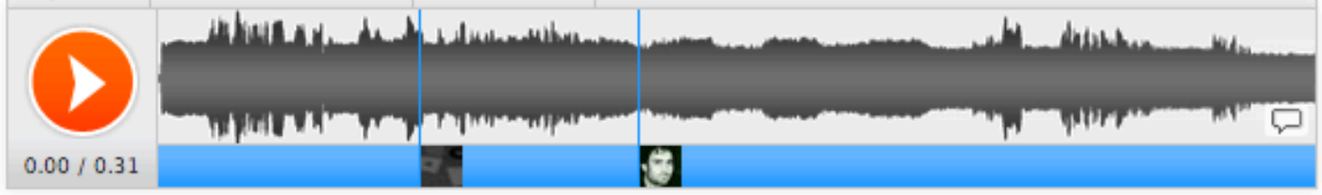
Share | Save to Favorites | Download | Radio

Eric ● **uploaded a new track**

melodic done drill at SoundCloud ReBase

Eric ● 1 day ago

Share | Save to Favorites | Download



Tanith **uploaded a new track**

Sommermix2012- Part 2

Tanith 1 day ago

257 ▶ | 3 💬 | 16 ♥ | 60 ⬇

Share | Save to Favorites | Download | Bass

Mini Update | More

	This Week	Total
Plays	15	608
Comments	-	44
Favoritings	-	11
Downloads	-	66
Profile Views	128	22509

Following 121 | Followers 1028 10+ | Groups 1

SHIGETO ★ ●

48 | 23110 — Following

Johannes Wagener ★ ●

63 | 1365 — Following

Kimbramusic ★ ●

18 | 4916 — Following

Hannes ★ ●

65 | 9751 — Following

Stream | You | Explore | Upload | Info

Your Stream

Dam Mantle ↻ seams — 15 hours
Brothers Fowl (LP preview mix)

+ | ↻ | ♥ | Share | ↓ | Buy | ▶ 239 | 💬 2 | ↻ 1 | ♥ 10

DW - Learn German — 8 hours
Bierselig | Sprachbar | Sep 19, 2012

+ | ↻ | ♥ | Share | ↓

2:54 Band ↻ Eric — 18 hours
2:54 - Killer

Play on!

Explore while listening to sounds non-stop with the new Continuous Play feature.

Who to follow — More

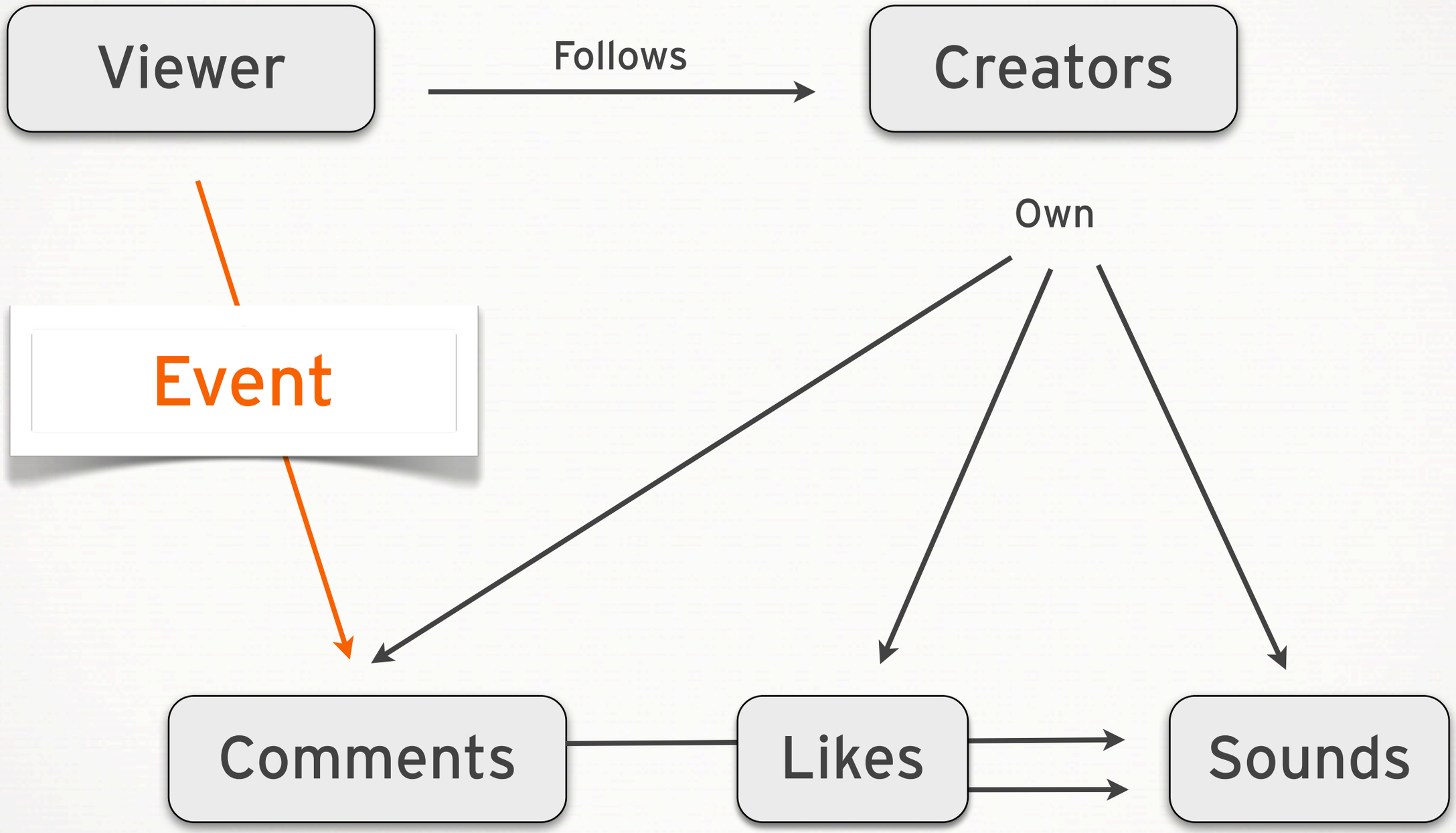
lauraherberg — 7 | 5942 — Follow

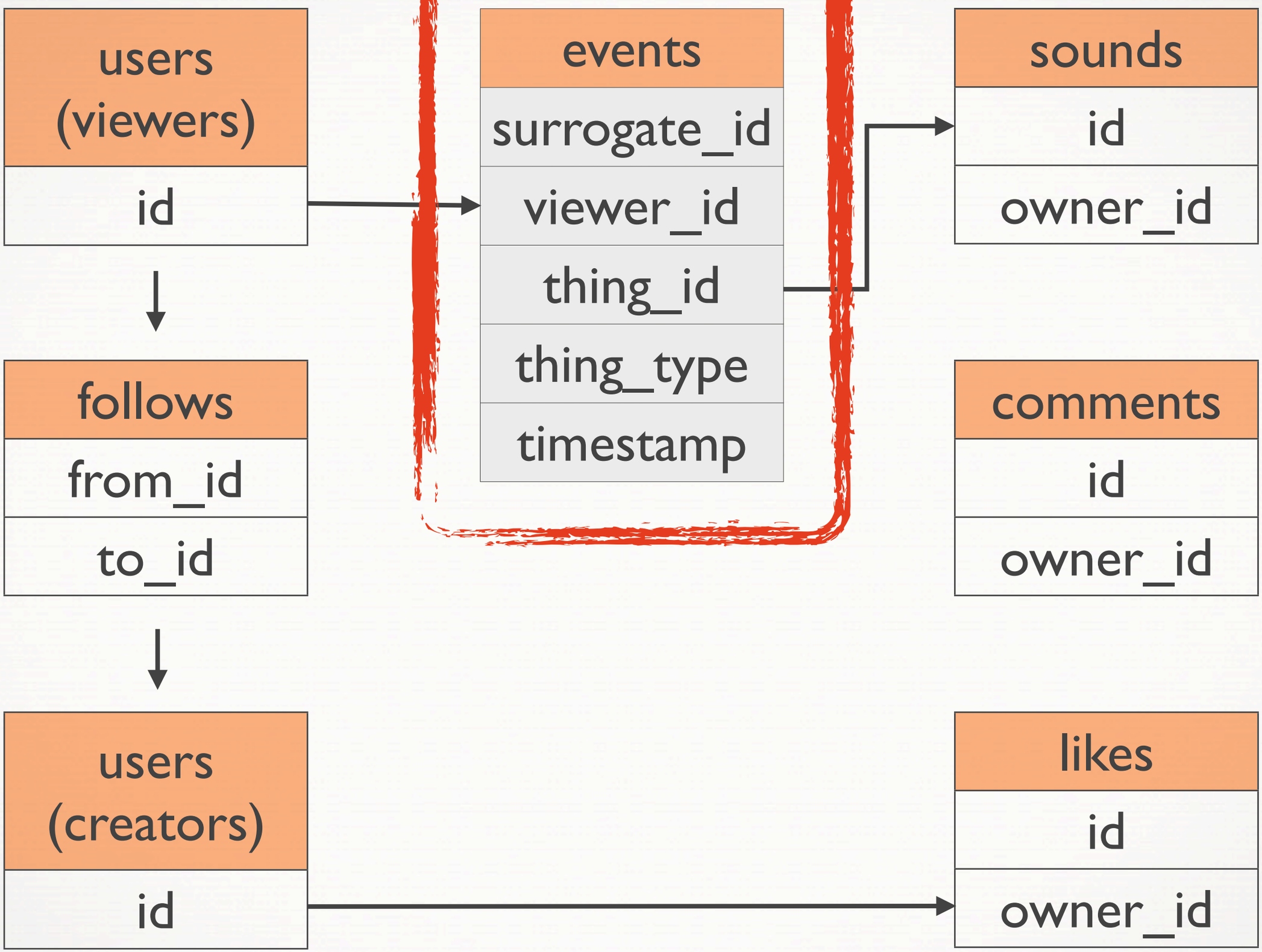
TheGaslightAnthem — 57 | 11754 — Follow

WorkingNow — 8 | 18309 — Follow

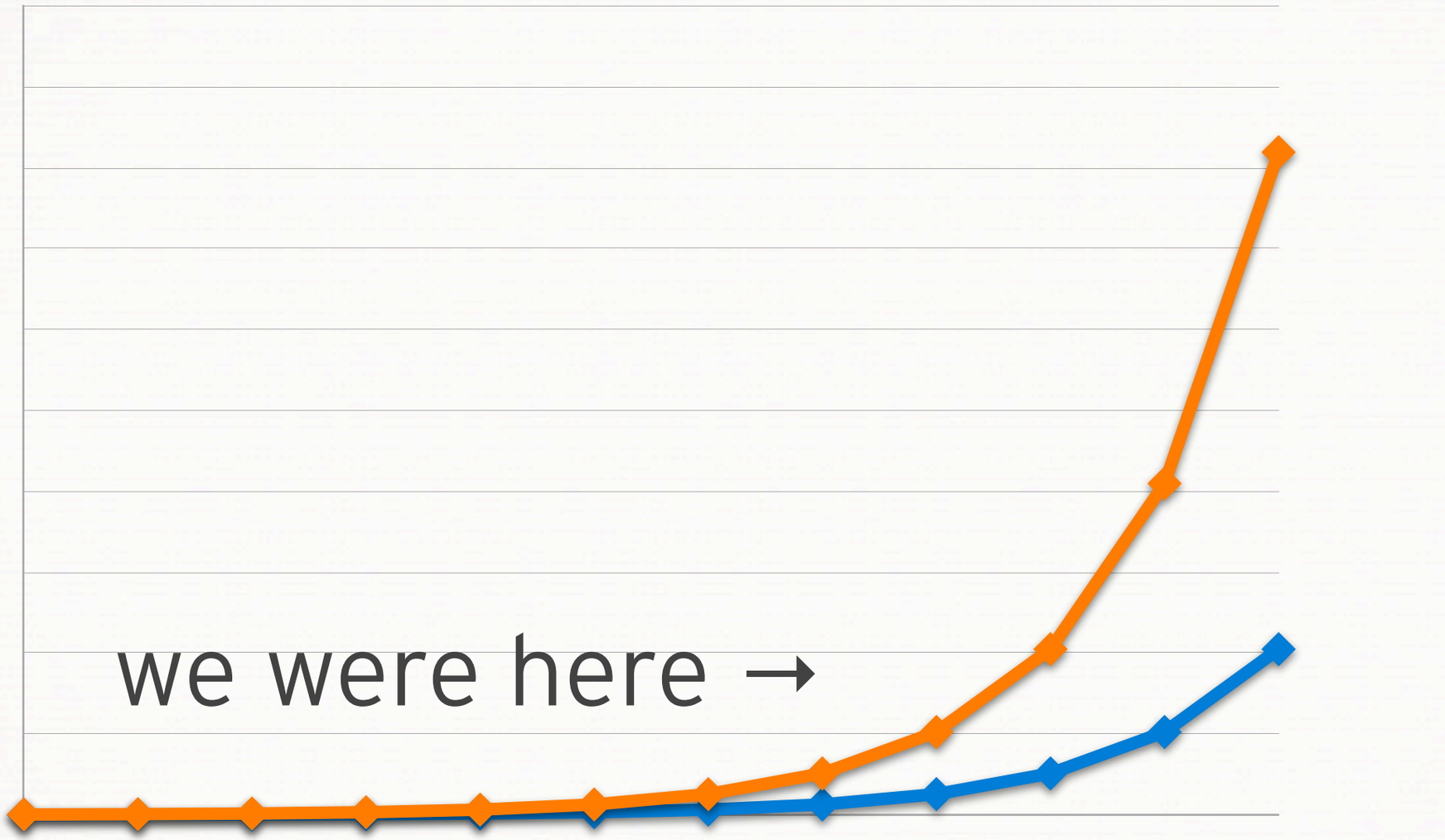
Recent activity — View all

C.A ELITE 2012 followed... — 11 hours
3 | 88 — Follow Back


Viewer
Follows
Creators
Own
Event
Comments
Likes
Sounds

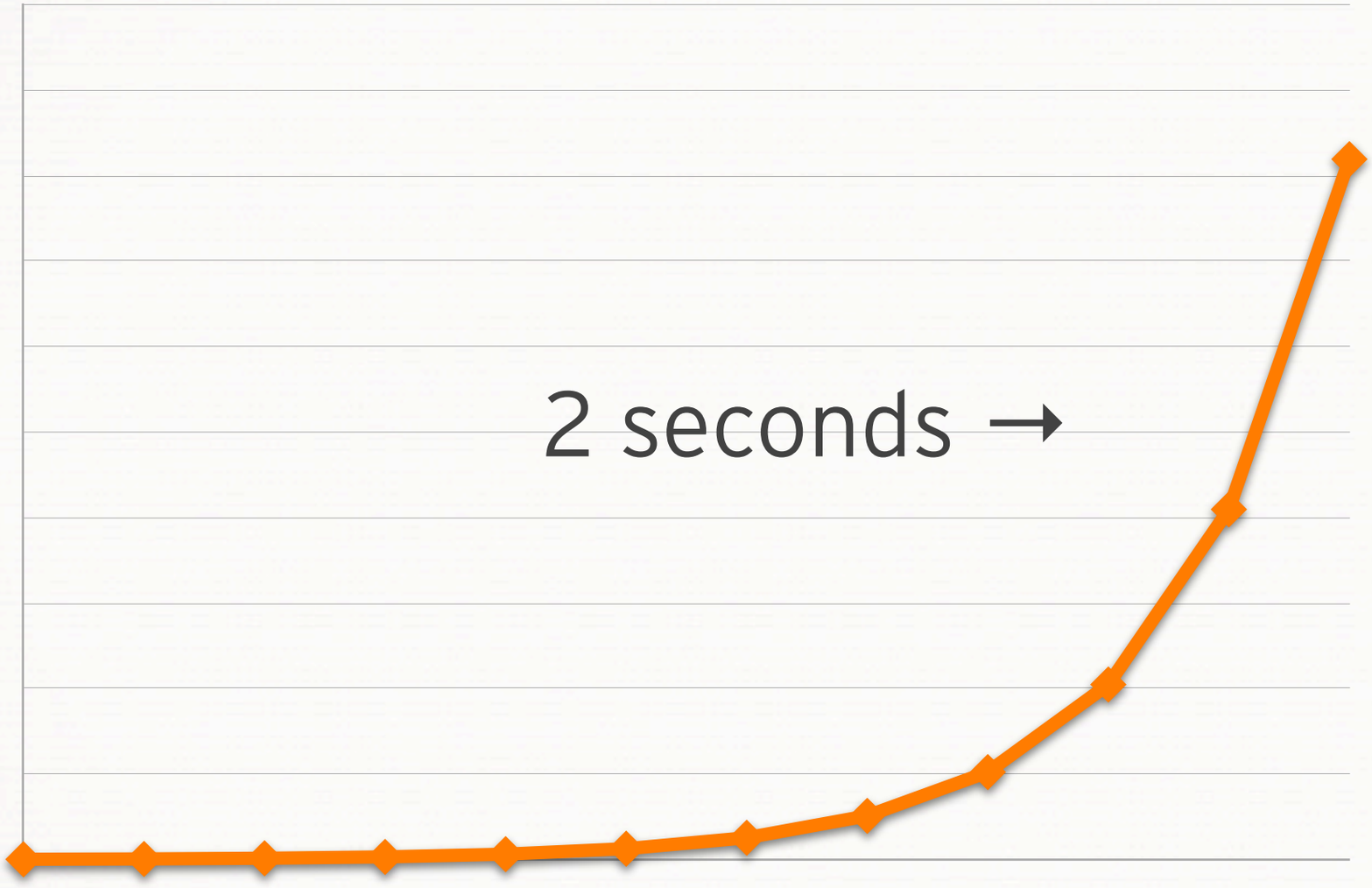

users
(viewers)
id
events
surrogate_id
viewer_id
thing_id
thing_type
timestamp
sounds
id
owner_id
follows
from_id
to_id
comments
id
owner_id
users
(creators)
id
likes
id
owner_id

GROWTH



Events **Registered Accounts**

LOAD TIME

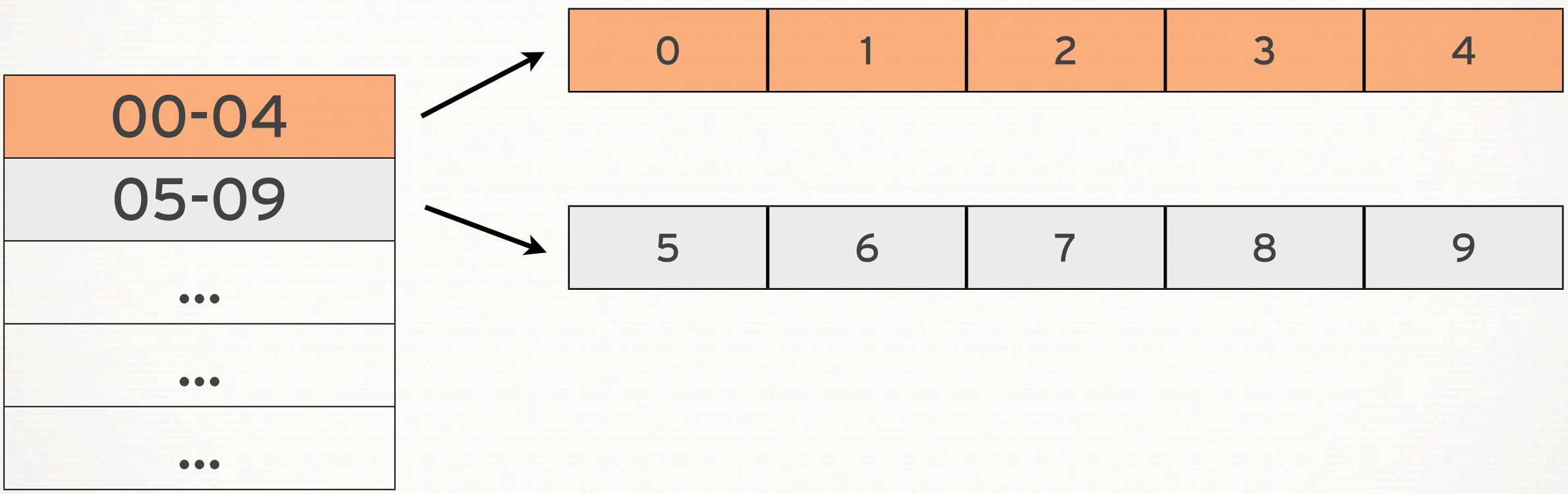


WHY

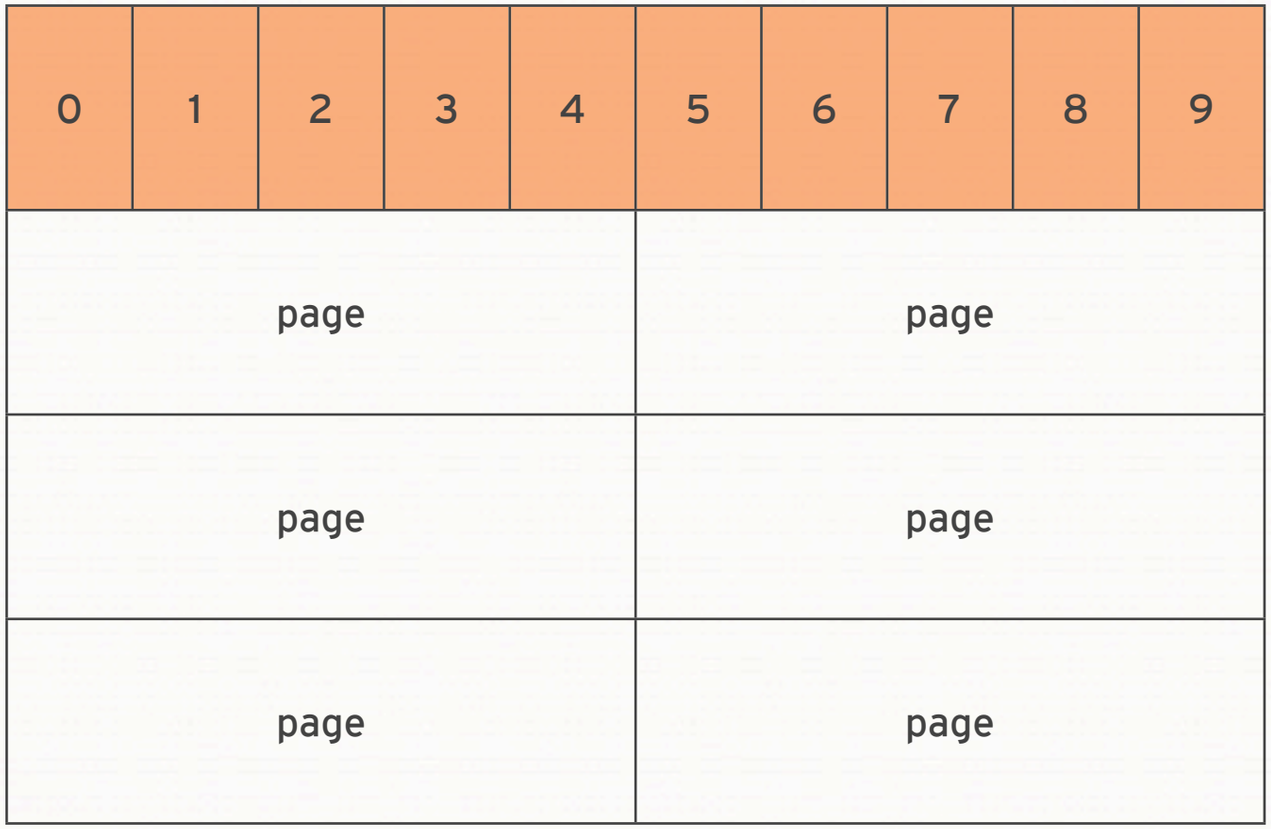
Time

id	viewer_id	thing_id
0	1234	…
1	1234	…
2	987654	…
3	1234	…
4	987654	…
5	8765	…
6	8765	…
7	987654	…
8	8765	…
9	8765	…
10	1234	…
…	…	…

B+ TREE



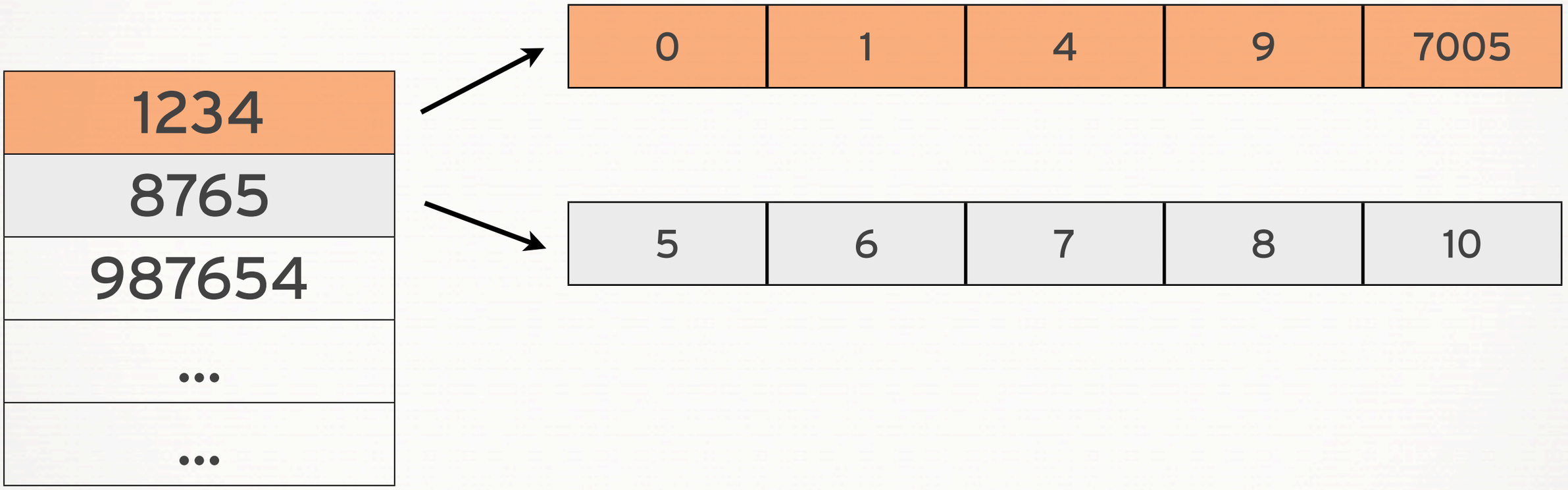
EVENTS ON DISK



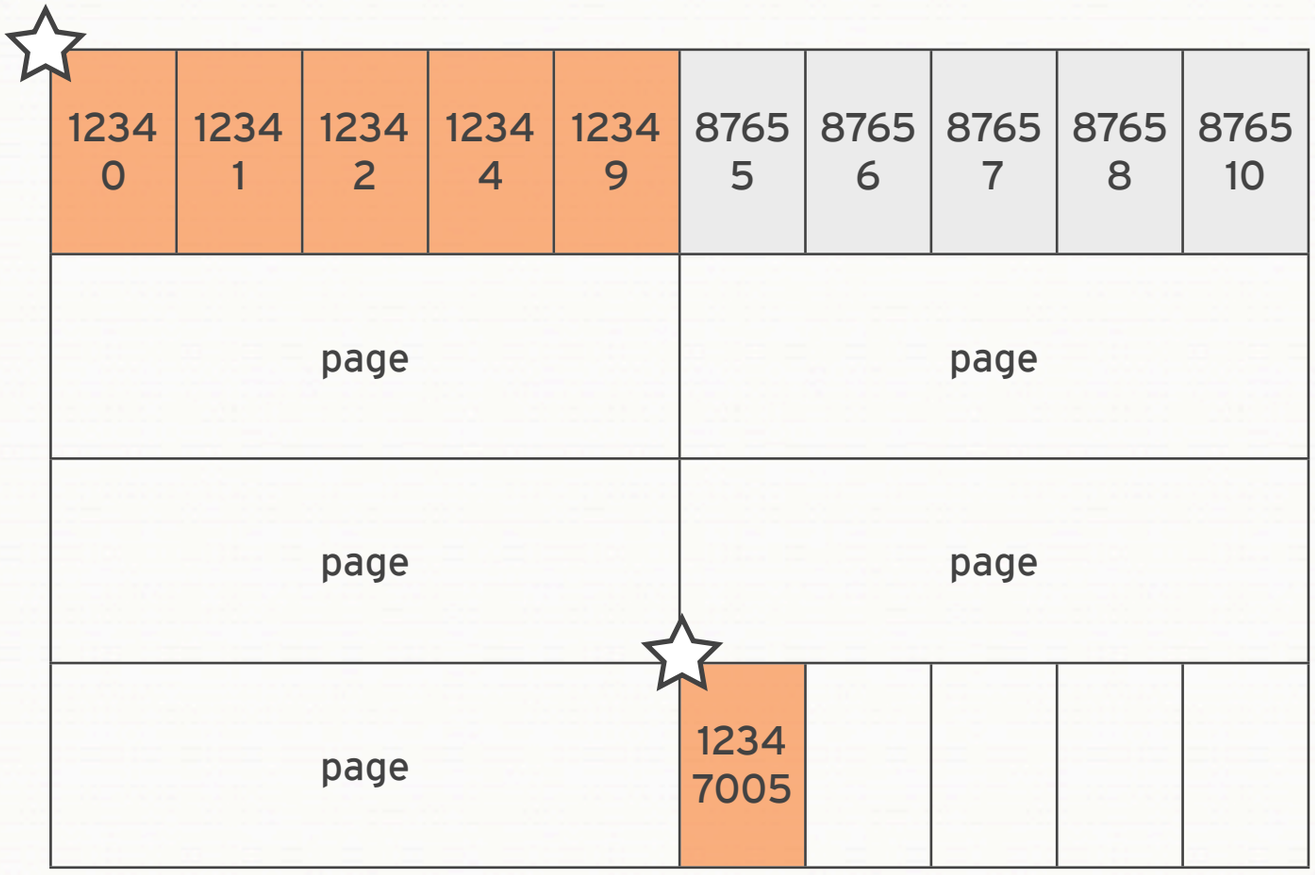
INDEXED

id	viewer_id	thing_id
0	1234	…
1	1234	…
2	987654	…
3	1234	…
4	987654	…
5	8765	…
6	8765	…
7	987654	…
8	8765	…
9	8765	…
…	…	…
7005	1234	…

INDEX

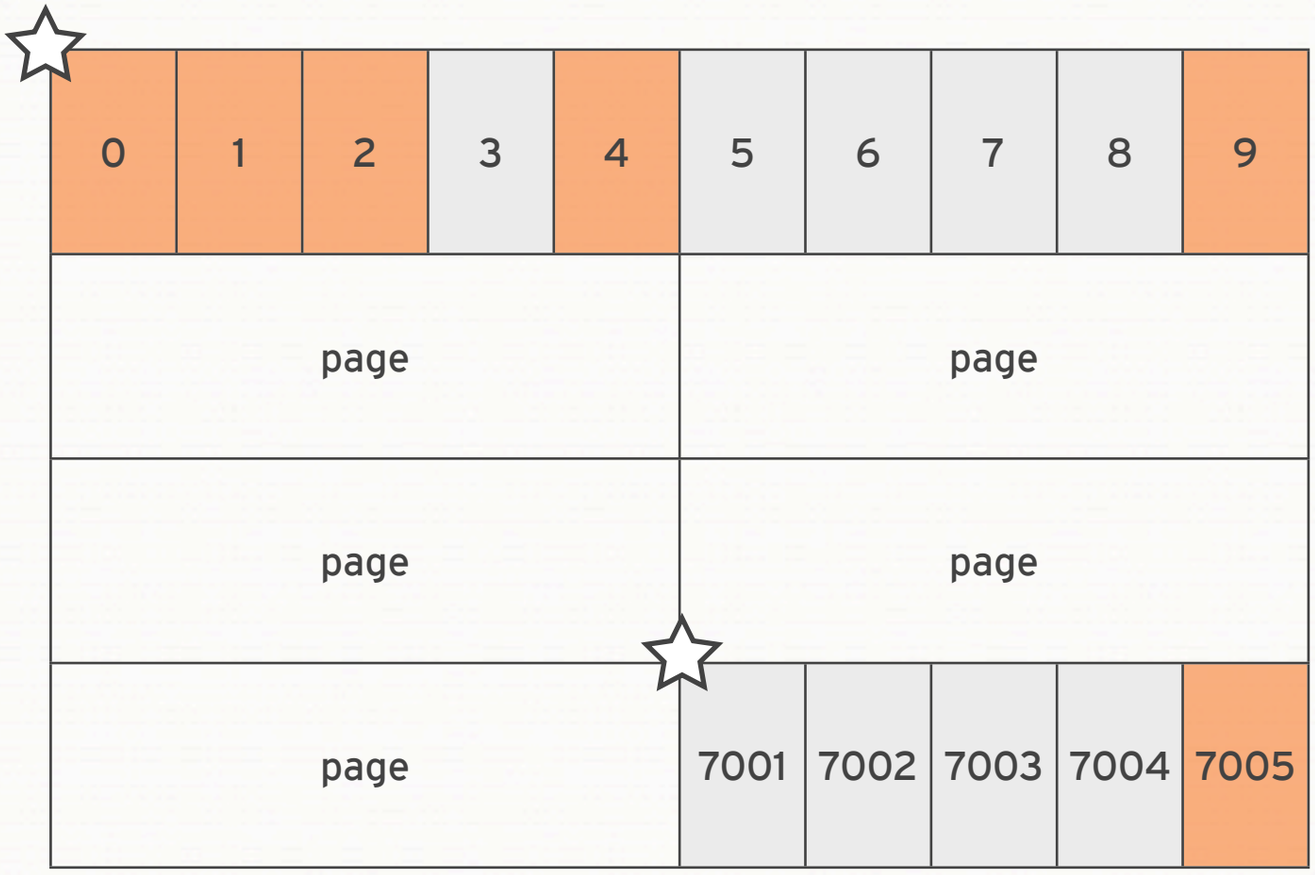


PAGE SPLIT



☆SEEKS ON READ

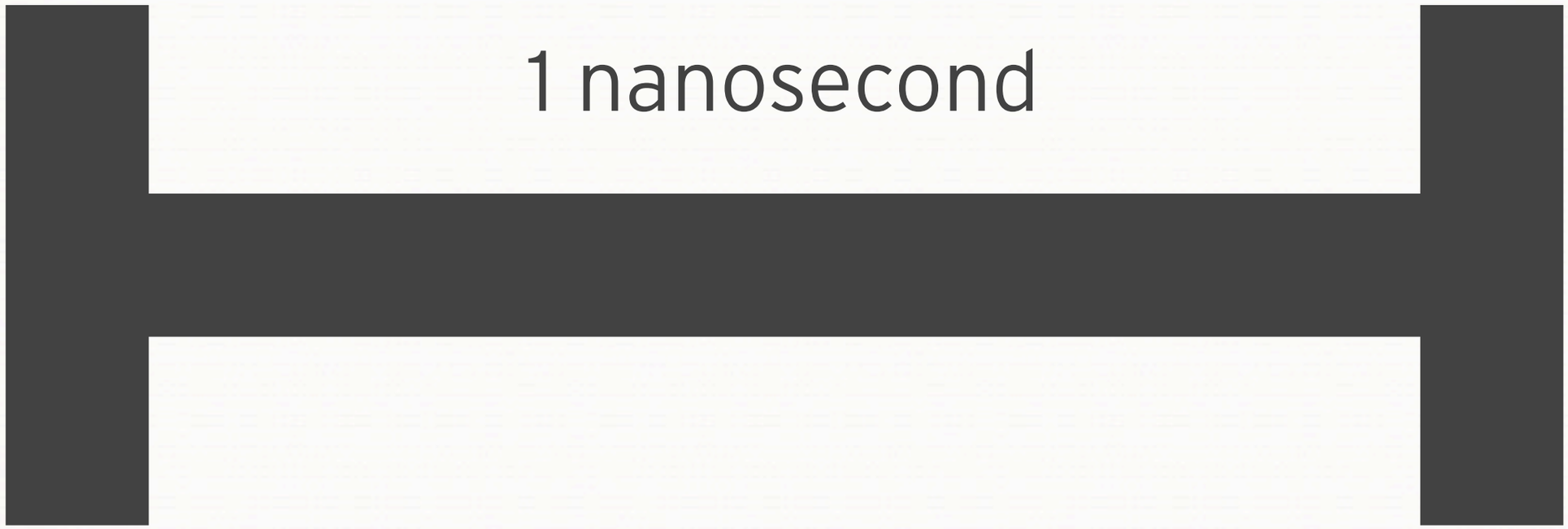
RECORD LOAD



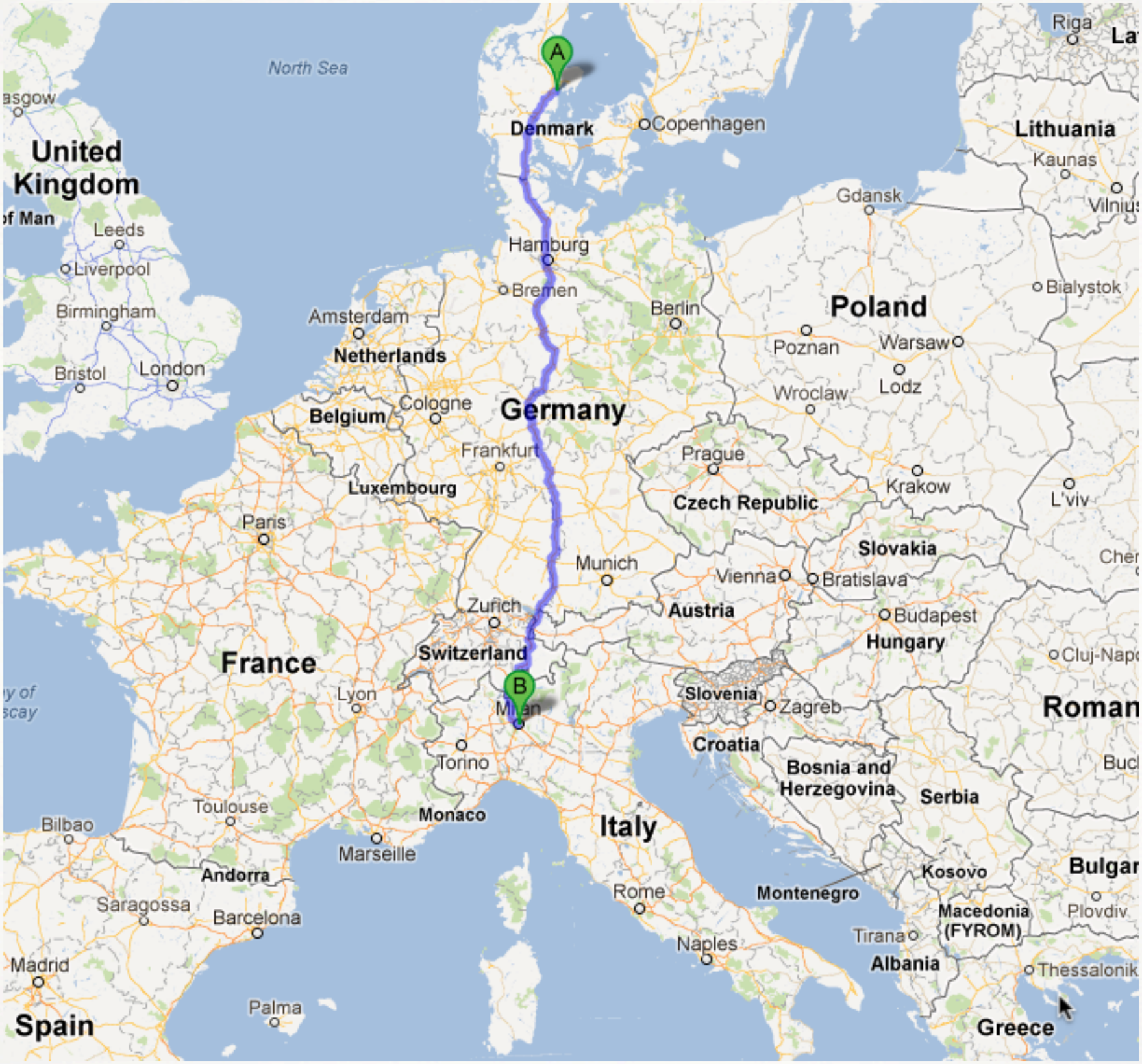
☆ SEEKS

SEEKS

1 nanosecond



10 milliseconds


North Sea
United Kingdom
Denmark
Copenhagen
Lithuania
Kaunas
Gdansk
Riga
Hamburg
Bremen
Berlin
Amsterdam
Netherlands
Poland
Poznan
Warsaw
Bialystok
Lodz
Wroclaw
Leeds
Liverpool
Birmingham
Bristol
London
Belgium
Cologne
Germany
Frankfurt
Prague
Czech Republic
Krakow
L'viv
Luxembourg
Paris
Munich
Slovakia
Vienna
Bratislava
Zurich
Austria
Budapest
Hungary
Switzerland
France
Lyon
Milan
Slovenia
Zagreb
Croatia
Bosnia and Herzegovina
Serbia
Torino
Monaco
Italy
Toulouse
Bilbao
Marseille
Andorra
Saragossa
Barcelona
Rome
Montenegro
Kosovo
Macedonia (FYROM)
Tirana
Albania
Naples
Madrid
Palma
Spain
Greece
Thessaloniki
Plovdiv
A
B

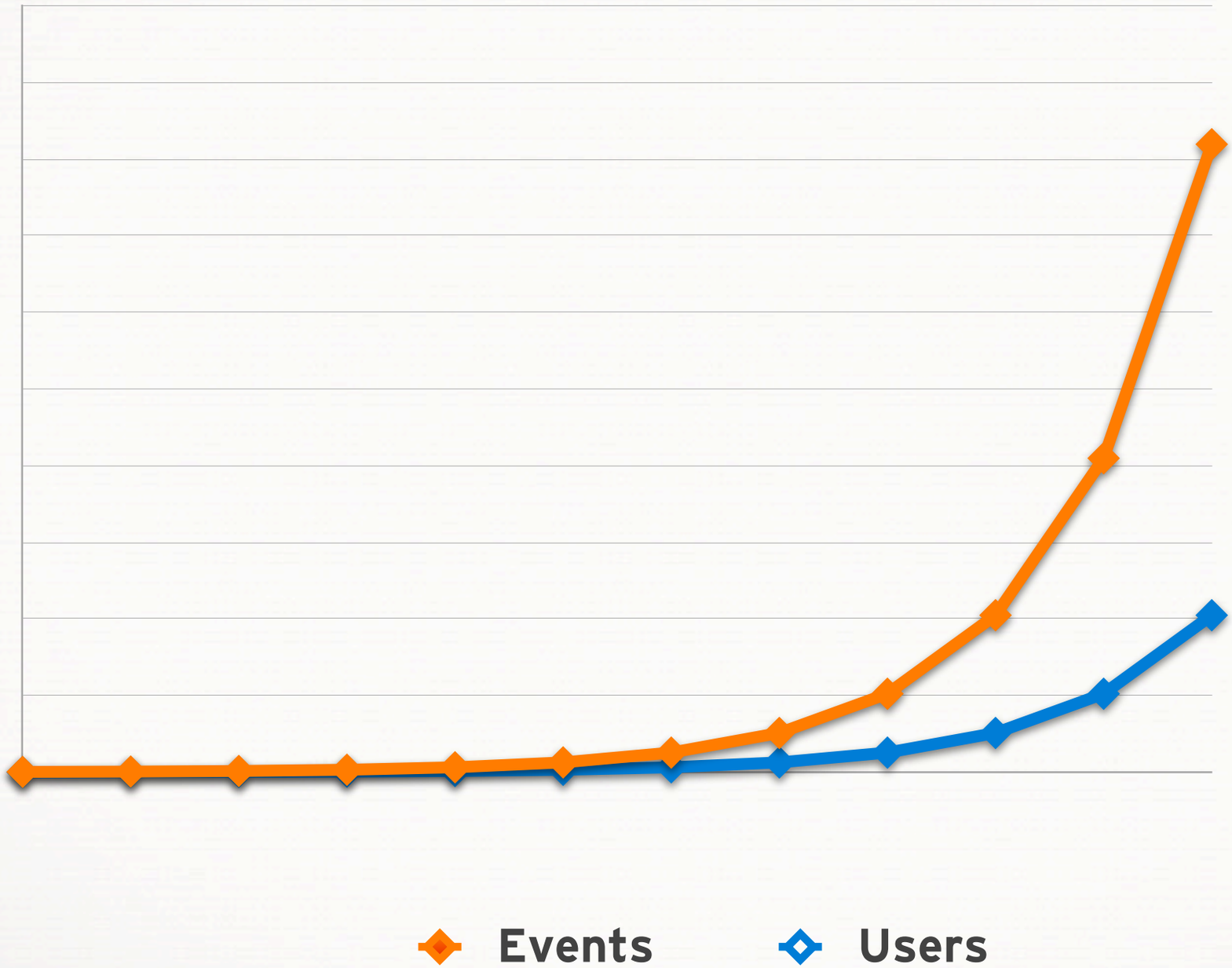
STORAGE

id	viewer_id	thing_id
0	1234	…
1	1234	…
2	987654	…
3	1234	…

SEEKS

765000	1234	…
…	…	…

GROWTH



10,000ft VIEW
(0.01ms)

EVENTUALLY CONSISTENT

HOT WRITES

COLD READS

PARTITIONED

KILL IT WITH FIRE

where fire = hard disks

EXTRACT
TRANSFORM
LOAD

WRITE TO MEMORY

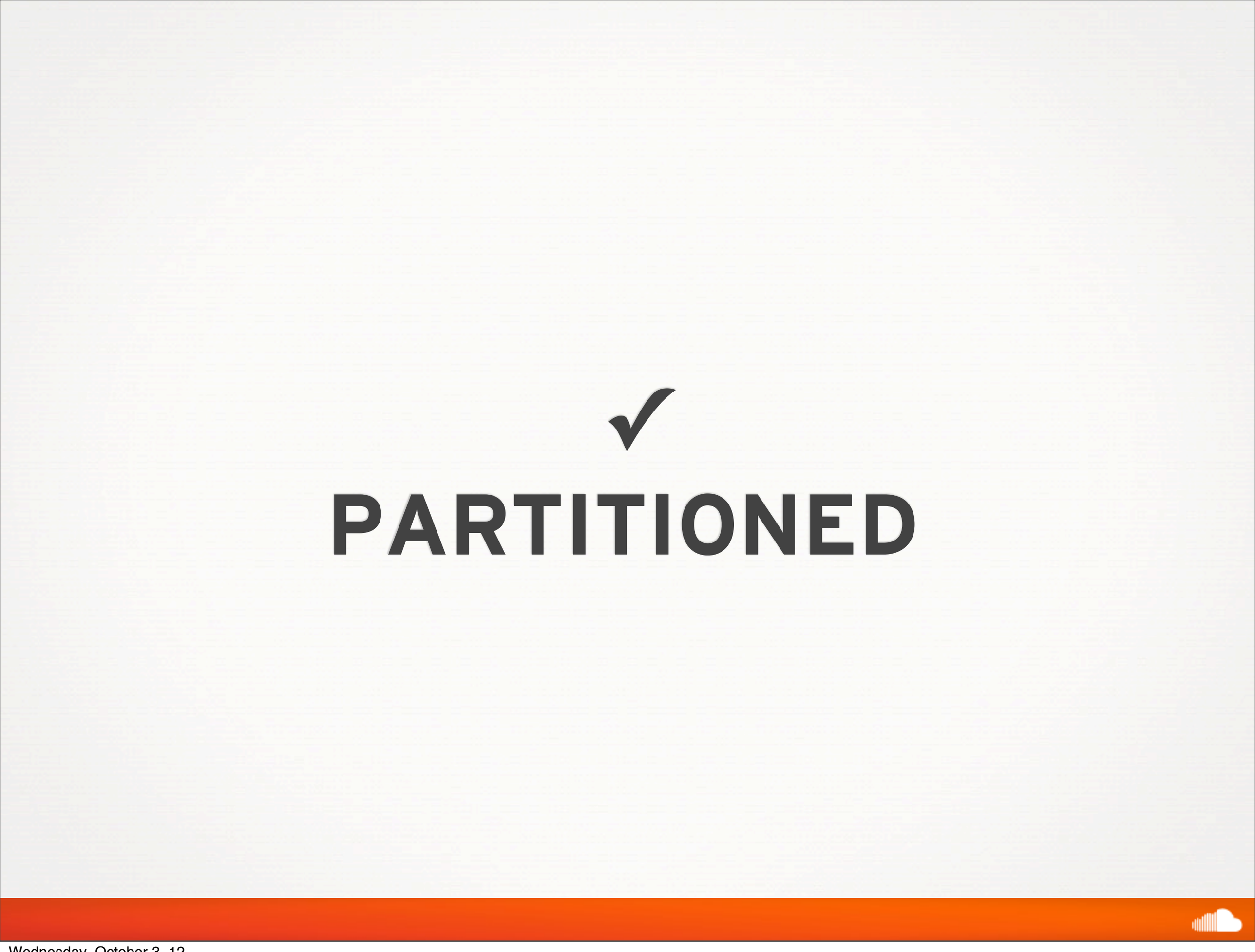
... magic happens ...

READ FROM DISK

CASSANDRA

✓

EVENTUALLY CONSISTENT


✓
PARTITIONED

✓

SEQUENTIAL WRITE

SEQUENTIAL READ

(on disk)

DATA MODEL

```
SELECT a, b, c
FROM table
WHERE table = ?
```

```
SELECT columns...
FROM table
WHERE table = ?
AND column > ?
```

```
SELECT columns...
FROM table
WHERE table = ?
AND column > ?
ORDER BY column
```

	C_1	C_2	C_3	C_n
R_1		V_1	V_2	
R_2	V_1		V_2	V_3
R_3	V_1	V_2	V_3	V_4

Row Key

Column Name

Value

	Name	Avatar	Email	Gender
1234		V_1	V_2	
8765	V_1		V_2	V_3
98765	V_1	V_2	V_3	V_4

Row Key
Column Name
Value

User ID
Property
Serialized Type

	1234	8764	98765	453627
1234		V_1	V_2	
8765	V_1		V_2	V_3
98765	V_1	V_2	V_3	V_4

Row Key
Column Name
Value

Left ID
Right ID
Edge Label

	1004	1003	1002	1001
1234		V_1	V_2	
8765	V_1		V_2	V_3
98765	V_1	V_2	V_3	V_4

Row Key
Column Name
Value

Viewer ID
Sequence
Thing ID

	t3-uuid	t2-uuid	t1-uuid2	t1-uuid1
1234		V_1	V_2	
8765	V_1			
98765				V_4

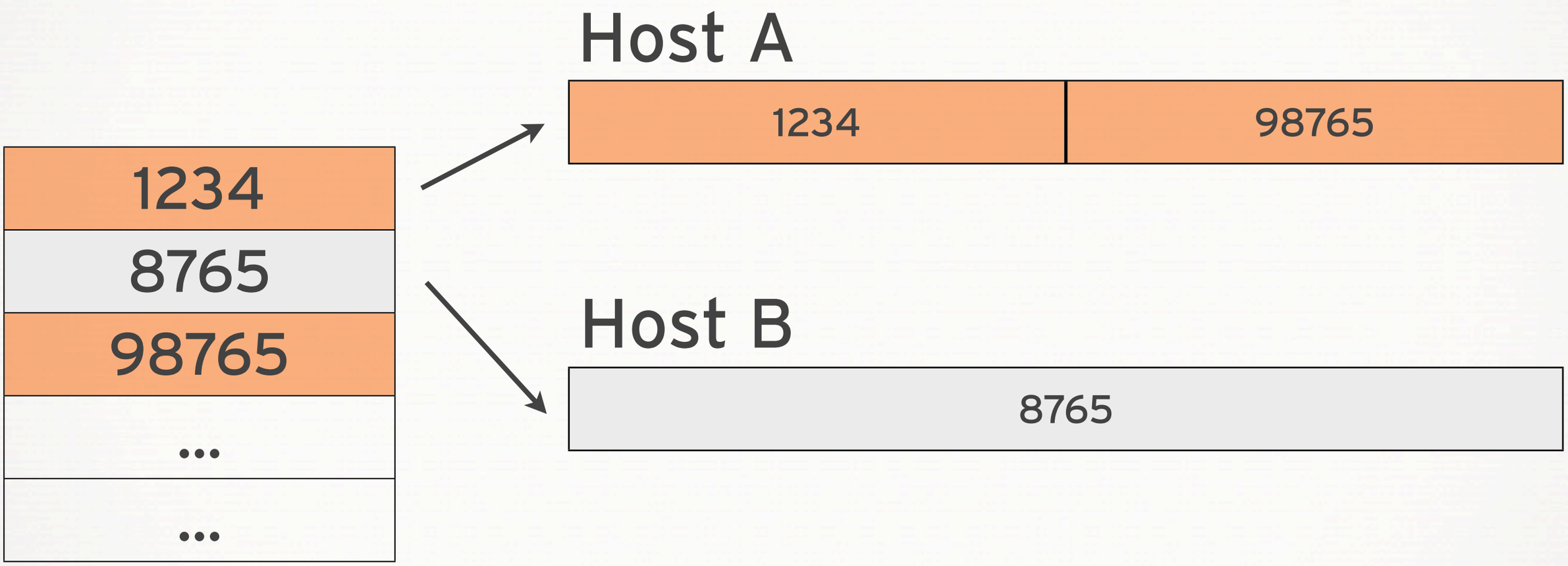
Row Key — Viewer ID

Column Name — TimeUUID (v1)

Value — Thing ID

STORAGE MODEL

RANDOM HOST PARTITIONING



MEMTABLE
(memory table)

SSTABLE
(sorted string table)

MEMTABLE

Time

1234	C_1	V_1
8765	C_2	V_2
98765	C_3	V_3
8765	C_4	V_4
8765	C_5	V_5
1234	C_6	V_6

SSTABLE

Sort
Row
Column

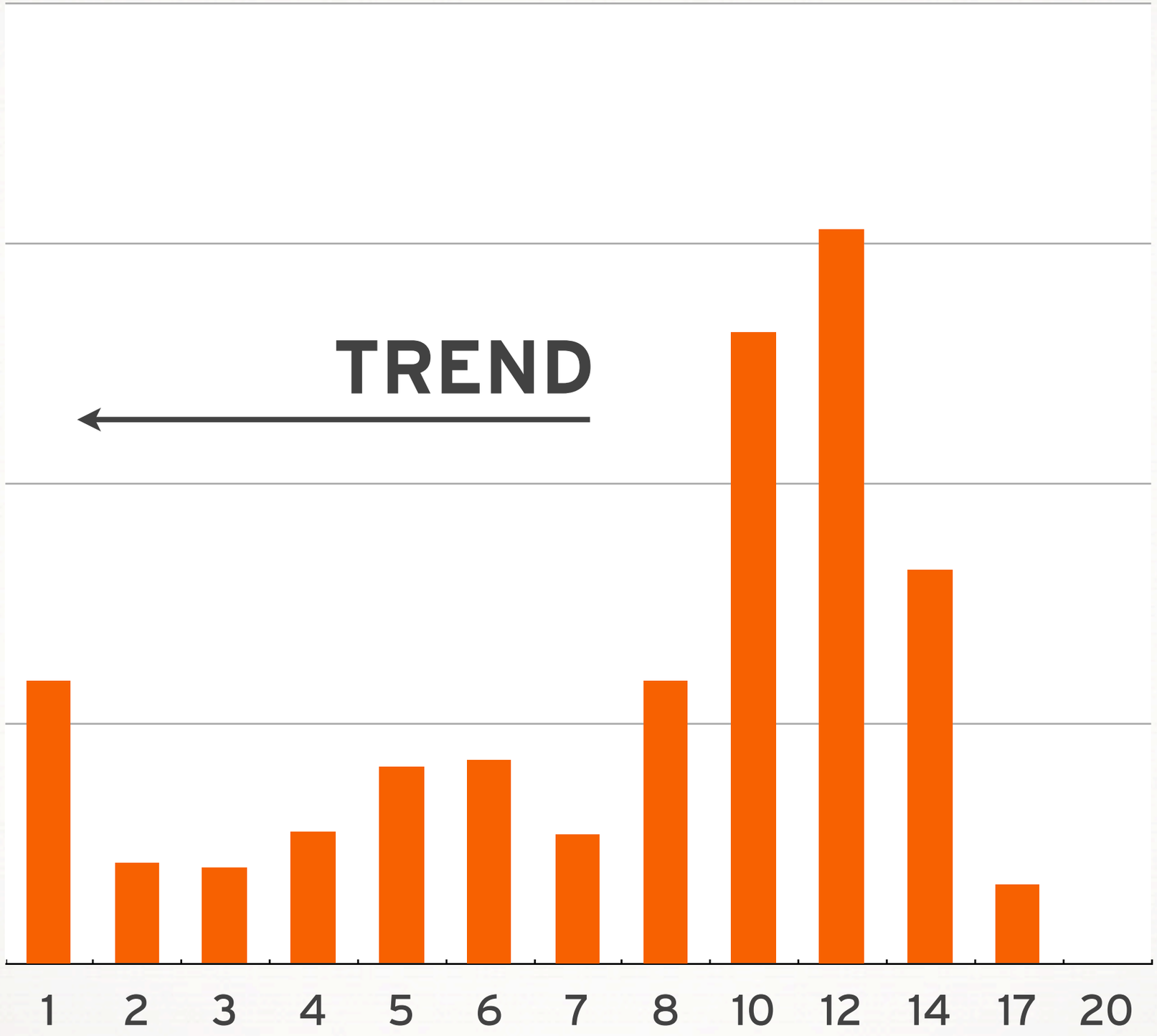
1234	C_1	V_1
1234	C_6	V_6
8765	C_2	V_2
8765	C_4	V_4
8765	C_5	V_5
98765	C_3	V_3

SS_1	1234	C_1	C_6	C_{11}	C_{13}
	8765	C_2	C_4	C_5	C_{12}

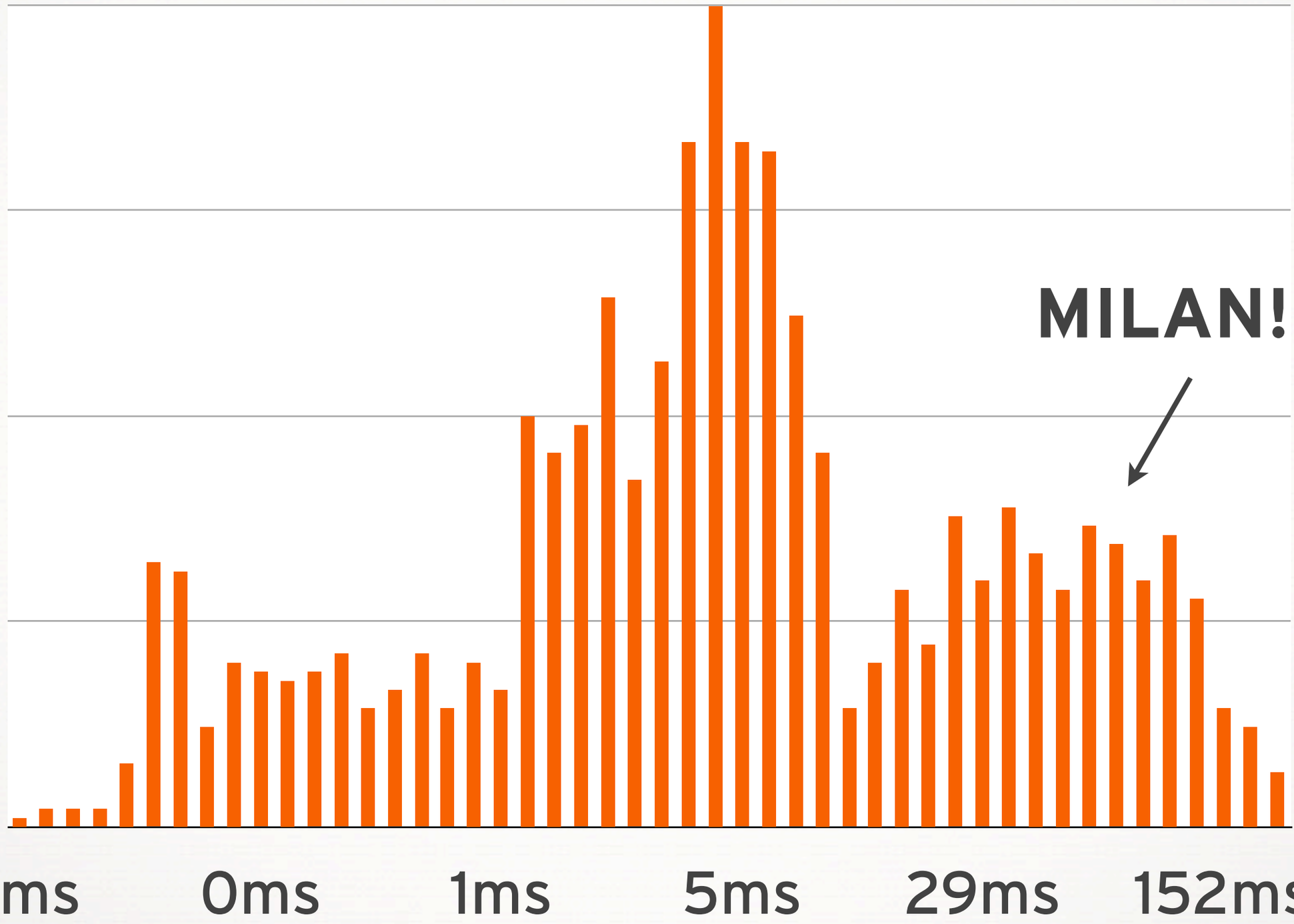
SS_2	1234	C_8	C_9	C_{17}	C_{19}
	98765	C_7	C_{18}	C_{20}	C_{21}

SS_3 (1+2)	1234	C_1	C_6	C_8	C_9
		C_{11}	C_{13}	C_{17}	C_{19}
	8765	C_2	C_4	C_5	C_{12}
	98765	C_7	C_{18}	C_{20}	C_{21}

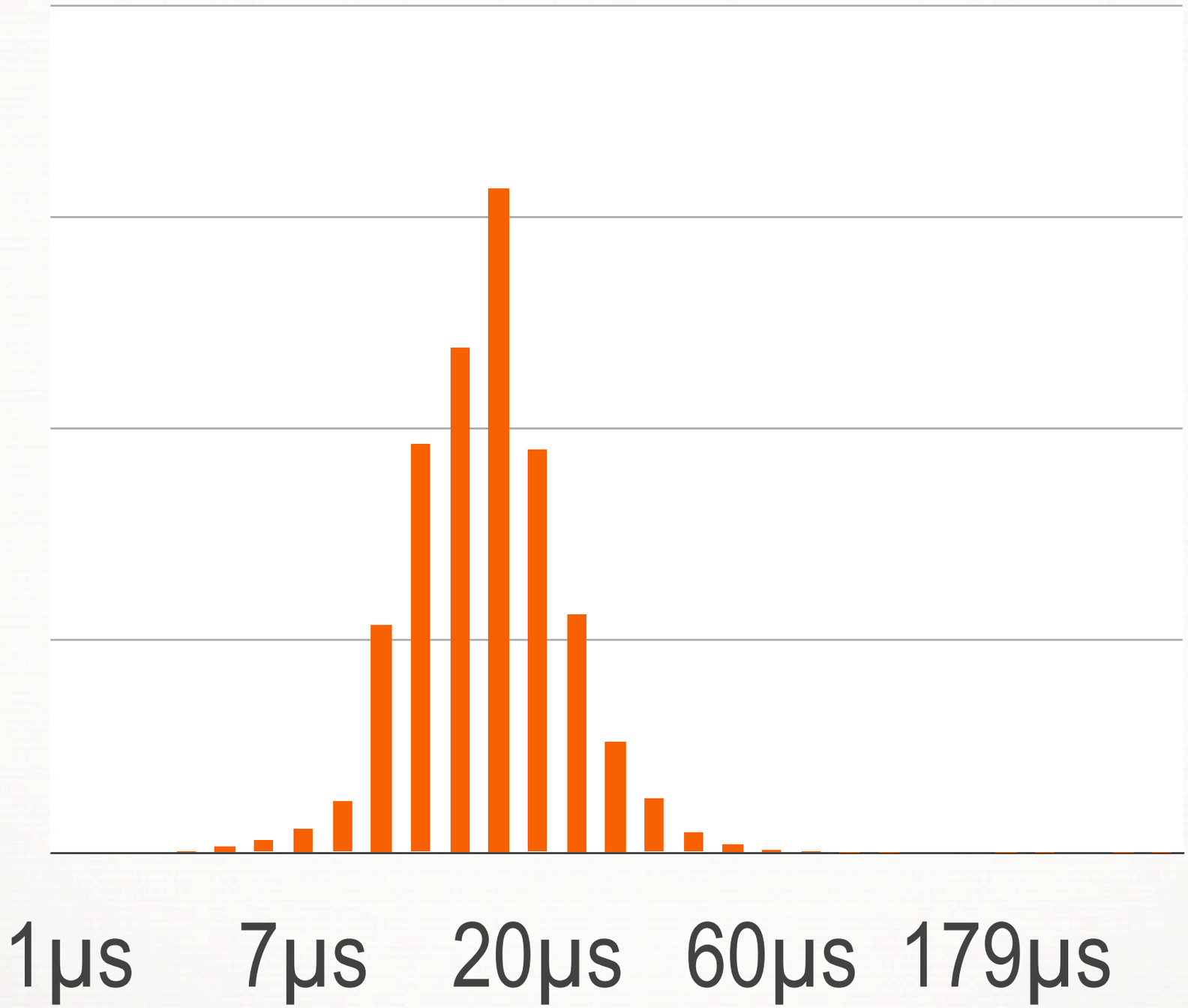
SSTABLES PER READ



READ LATENCY (milliseconds)



WRITE LATENCY (microseconds)



CASSANDRA

MYSQL + INNODB

SPINNING DISKS

Sean Treadway

http://github.com/streadway

FITTING THE MODEL

WITH YOUR MATERIALS