

twitter 

UC BERKELEY - CS10



@raffi
raffi

Giving a talk at Cal's CS10 - The
Beauty and Joy of Computing.
inst.eecs.berkeley.edu/~cs10/fa10

20 Oct via [Twitter for iPhone](#) ☆ Favorite ↩ Reply 🗑 Delete

from UC Berkeley
155 Donner Lab
Berkeley, CA
[View Tweets at this place](#)

TO START... WHAT'S A TWEET?

FOLLOW ME!

twitter.com/raffi
or send a SMS to 40404
follow raffi





James Buck

@jamesbuck Oakland, CA, USA

US-based photojournalist and multimedia producer.

<http://jameskarlbuck.com>

“Rudimentary communication
among individuals in real time
allows many to move together as
one” - @biz

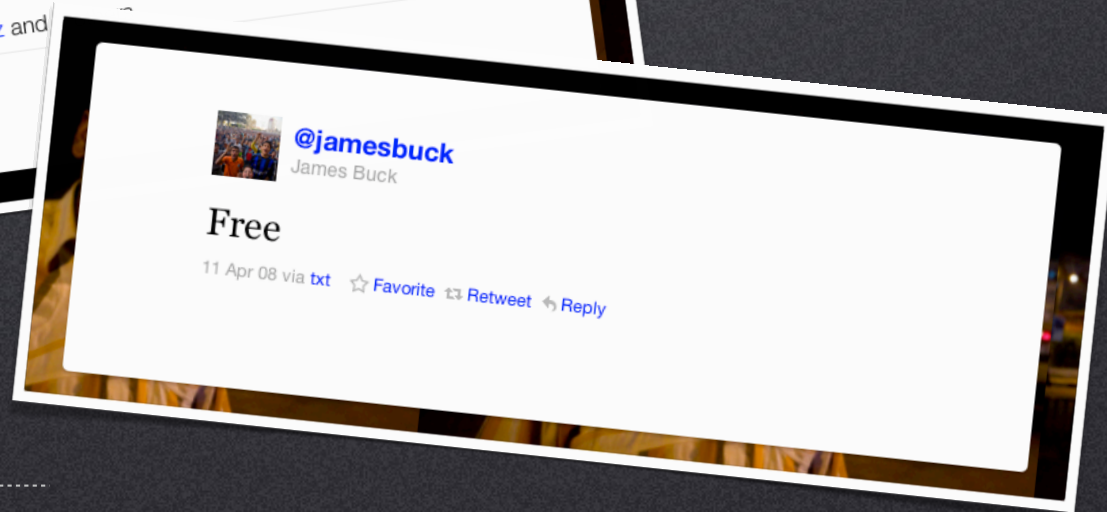


ONE WORD

April 10, 2008
4:33:23 PM

ONE WORD

April 10, 2008
4:33:23 PM



AND IT'S OVER

April 11, 2008
1:27:26 PM

HAITI

January 2010

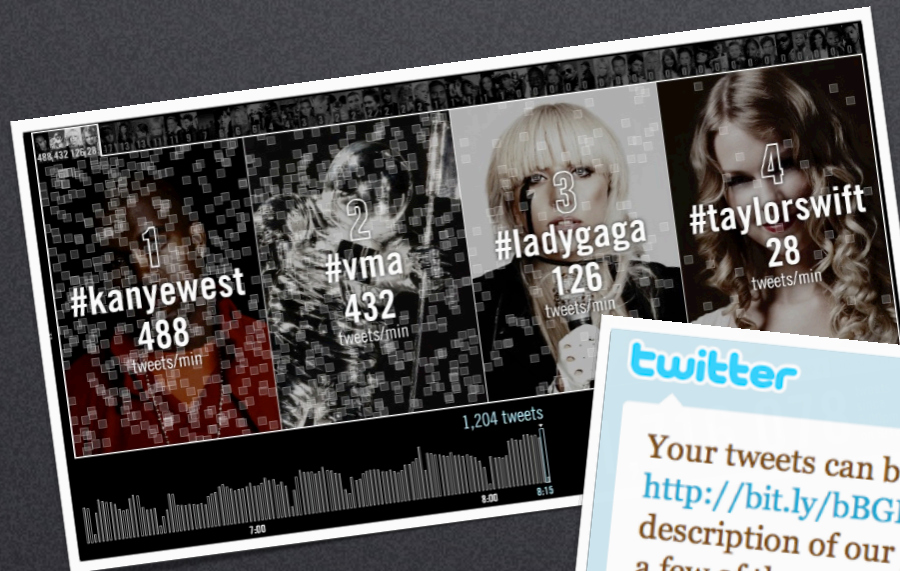


CHILE

February 2010

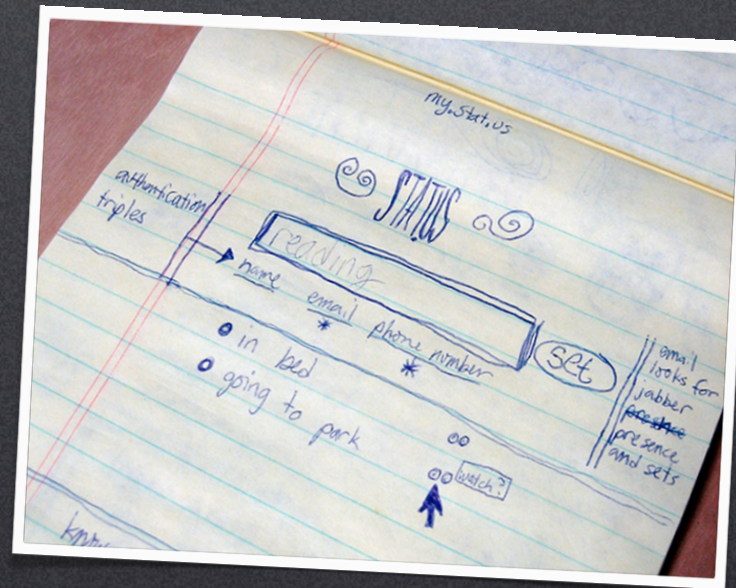
MTV VMAs

September 2010



EMMYS

August 2010



The humble beginnings

HOW DOES TWITTER WORK?

**TWITTER IS JUST A “STRUCT”, A GRAPH,
& A LIST**

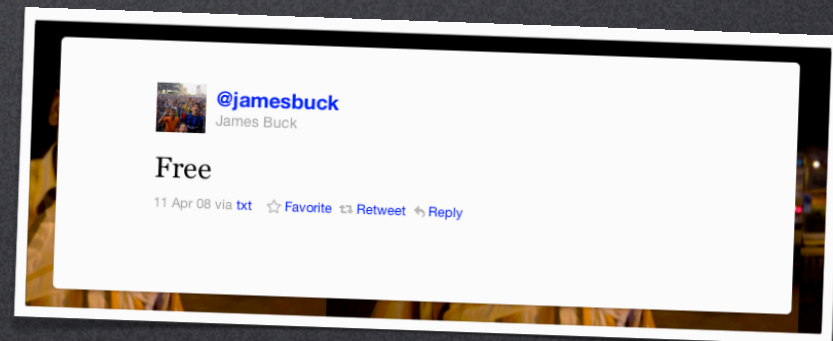
A TWEET IS A SIMPLE THING



140 characters



WHAT'S THE DATA STRUCTURE OF A TWEET?



```
{  
  "id" : 787167620,  
  "text" : "Free",  
  "date" : "2008-04-11 13:27:26"  
  "user" : "jamesbuck"  
}
```

*not quite, but something like this

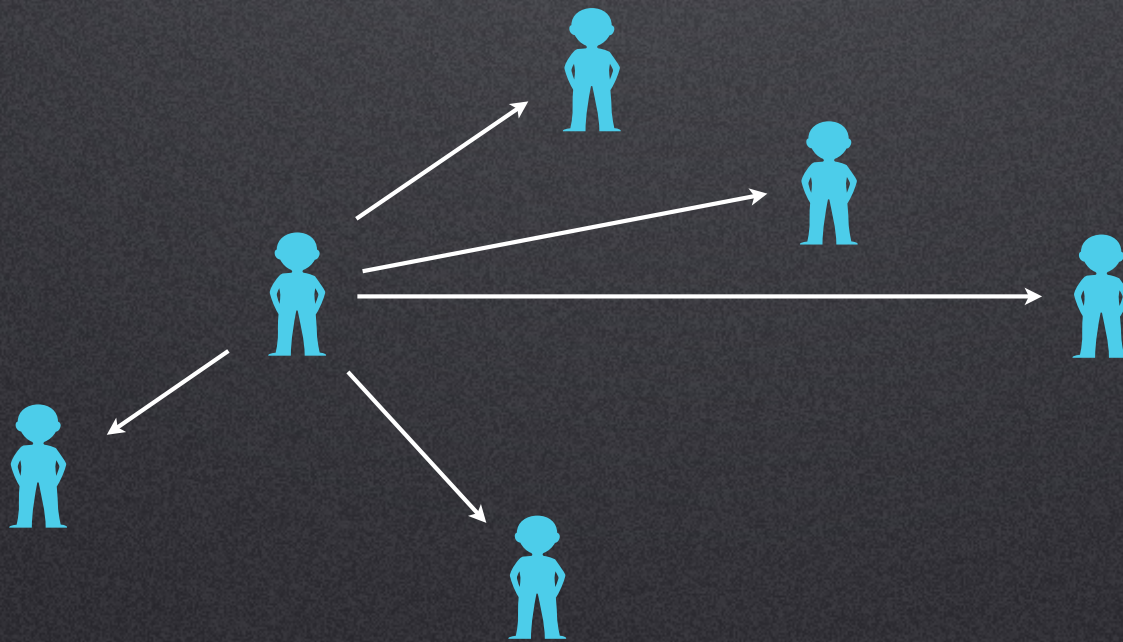
WHERE DO THEY GO?



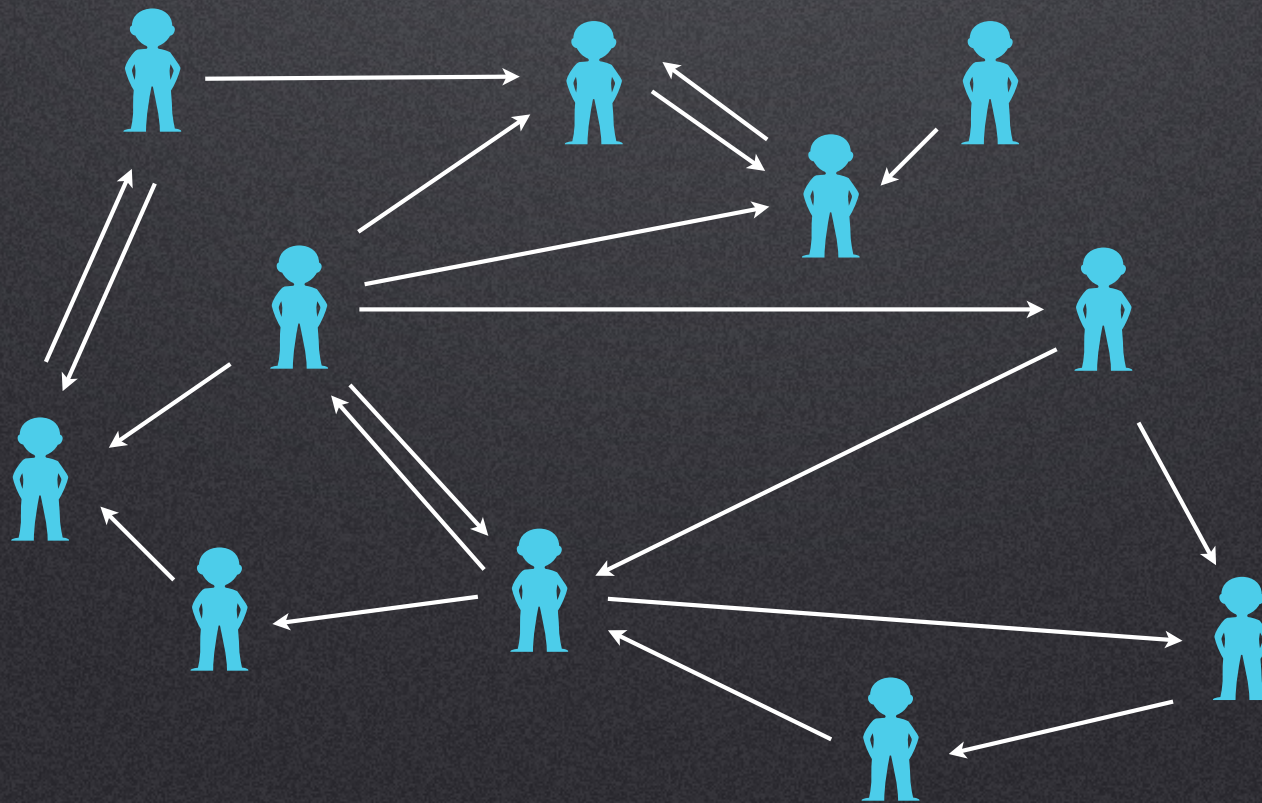
Following
Followed by



TWEETS MULTIPLY!




IT'S A HUGE GRAPH




VIEWING TWEETS


Timeline @Mentions Retweets Searches Lists


1 new tweet

 **sandofsky** Ben Sandofsky
"Clean All Targets" should be renamed "Blow on Cartridge."
from Castro, San Francisco
3 minutes ago

 **pandemona** Sara Haider
not the kind of day I expected it to be.
3 minutes ago

 **jennadawn** Jenna
Mumford & Sons tomorrow @thearfield. Yeehaw!
4 minutes ago

 **jess** Jessica Verrilli
@keltonlynn just say the word. I use to have a buncha sticks in the office.
5 minutes ago

 **ashedryden** Ashe Dryden  by abraham
@Abraham "You have been betrayed by the space pirates. You should have known better." +1
from Historic Mitchell Street, Milwaukee

HOME TIMELINE

Timeline @Mentions Retweets Searches Lists

1 new tweet

sandofsky Ben Sandofsky
"Clean All Targets" should be renamed "Blow on Cartridge."
from Castro, San Francisco
3 minutes ago

pandemona Sara Haider
not the kind of day I expected it to be.
3 minutes ago

jennadawn jenna
Mumford & Sons tomorrow @thearfield. Yeehaw!
4 minutes ago

jess Jessica Verrilli
@keltonlynn just say the word. I use to have a buncha sticks in the office.
5 minutes ago

ashedryden Ashe Dryden ↻ by abraham
@Abraham "You have been betrayed by the space pirates. You should have known better." +1
from Historic Mitchell Street, Milwaukee

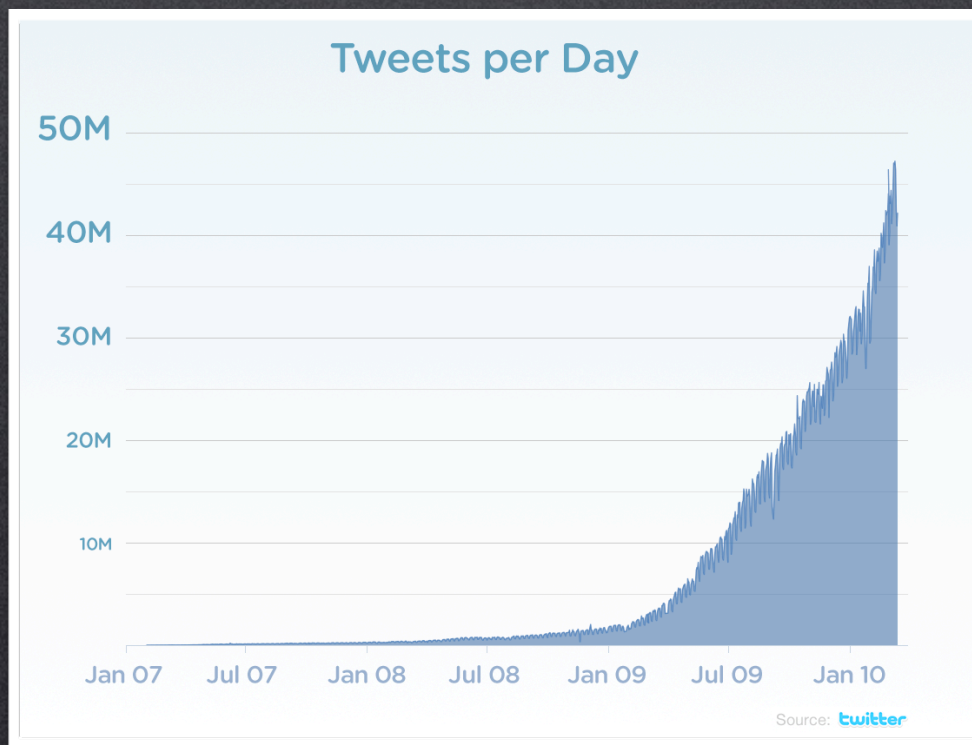
[27886578700,
27856404785,
27727644786,
27679412331,
27594550593,
27583976236,
27580246968,
27576503025,
27545926564,
27545450496,
27539864336...

IS THAT IT? ... YES! ... SORTA ...

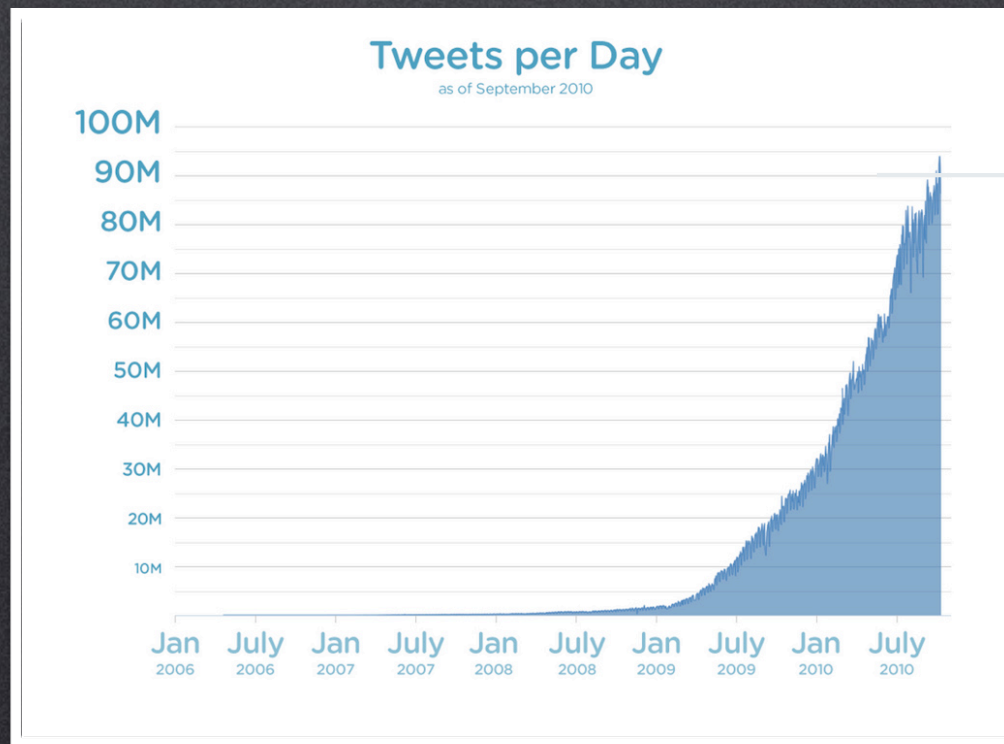
THINGS GET MORE COMPLICATED AT SCALE

AND SCALING REAL-TIME IS HARD

HOW MANY TWEETS ARE THERE?



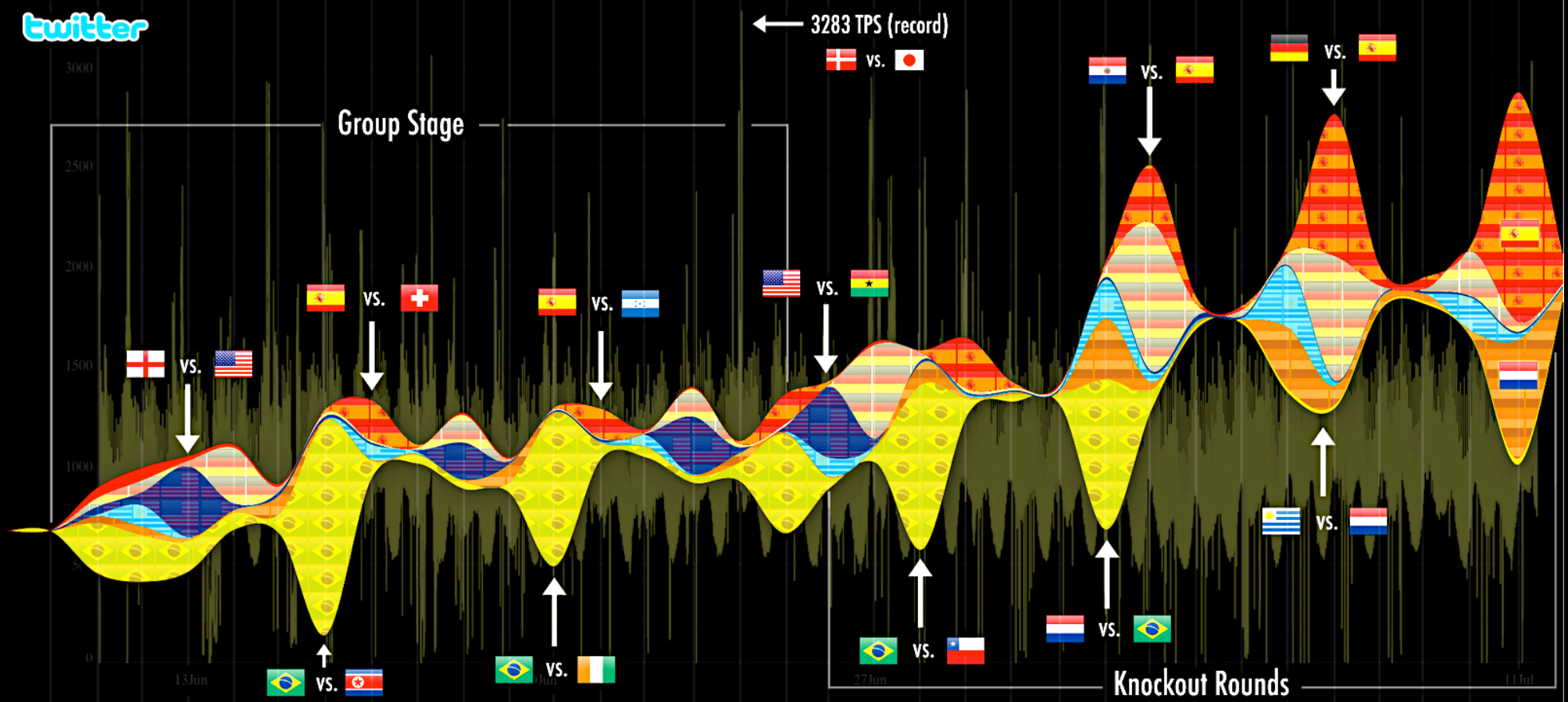
HOW MANY TWEETS ARE THERE?



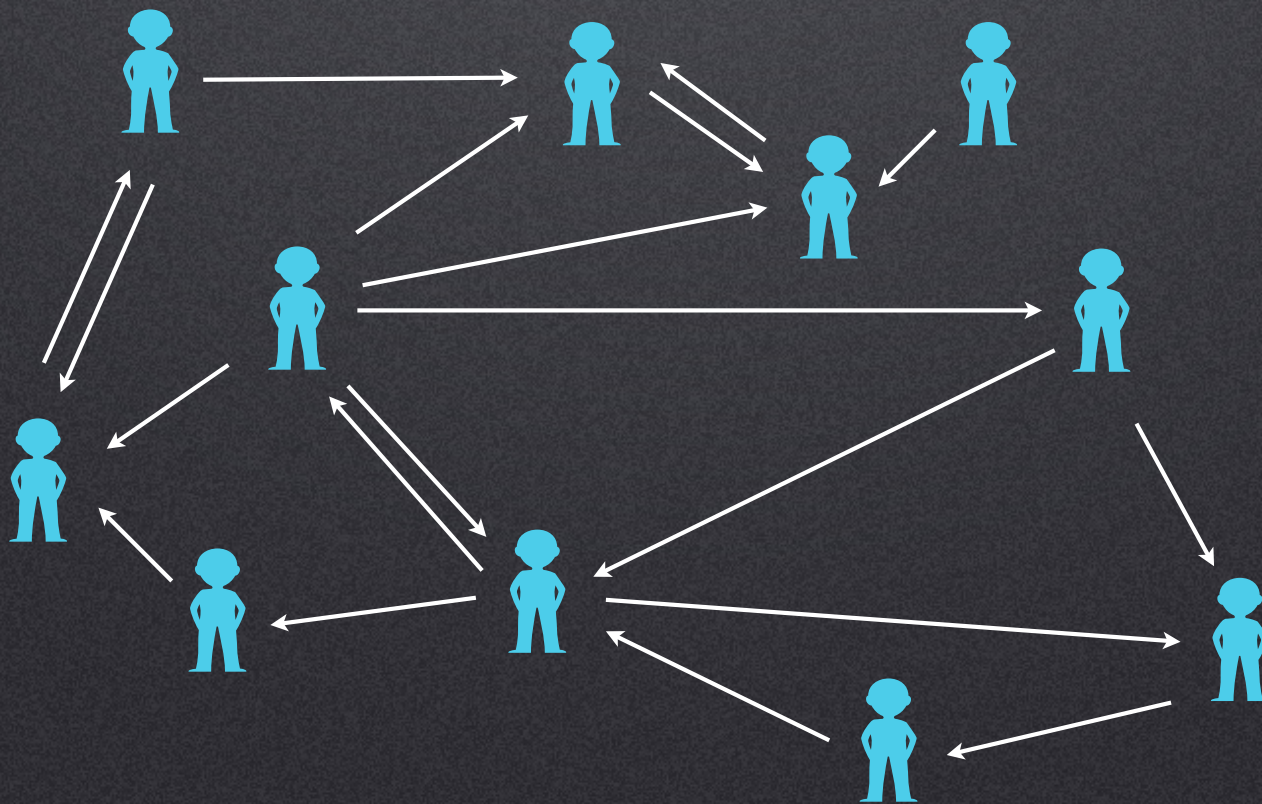
90M!

90M tweets
per day \approx 1000 tweets
per second

twitter



LOTS OF DELIVERIES



TWEETS MULTIPLY

Your Tweets **2,963**

 5 hours ago: OH @wil - "@raffi's car is like a gentoo package, it emerges onto the freeway."

Following **498**



Favorites **36**

★ **nelson** Awesome dinner at 10 Tables with my uberUROP @raffi. We've moved fr...

Followers **5,162**

Listed **447**

☰ Recently listed in: [platser](#), [SocialMediaPeople9](#), [Tech](#), [Twitter-Team](#), [GenteBoa](#)

5,162 people
need to see my
tweet



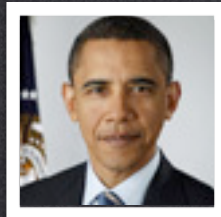
@ladygaga

mother monst[†]er
6.8 million followers



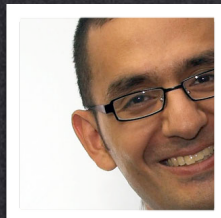
@justinbieber

Justin Bieber
5.7 million followers



@BarakObama

44th President of the United States
5.6 million followers



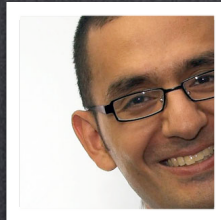
@raffi

me!
5.1 thousand followers



@ladygaga

mother monst[†]er
6.8 x 10⁶ followers



@raffi

me!
5.1 x 10³ followers

differs by 10³



@ladygaga

mother monster
6.8 x 10⁶ followers

scaling 10⁶

10^6
deliveries
per minute = 6000
deliveries
per
millisecond

**A SYSTEM IS ONLY AS STRONG AS ITS
WEAKEST LINK**



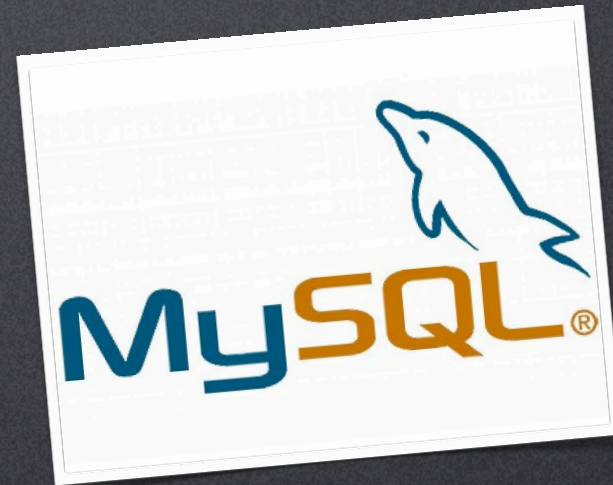
80 MB
per sec

10 ms
per seek

10 ms = 100 seeks
per seek = per second

MySQL

Can't generate IDs fast enough
Centralized and a single point of failure



snowflake

Highly available and uncoordinated (10kqps)
Compatible with the **twitter**  ecosystem
<http://github.com/twitter/snowflake>

HOW BIG ARE THEY?

1 tweet text = 140 characters
≈ 200 bytes

1000 tweets
per second



195 KB/sec



11 MB/min

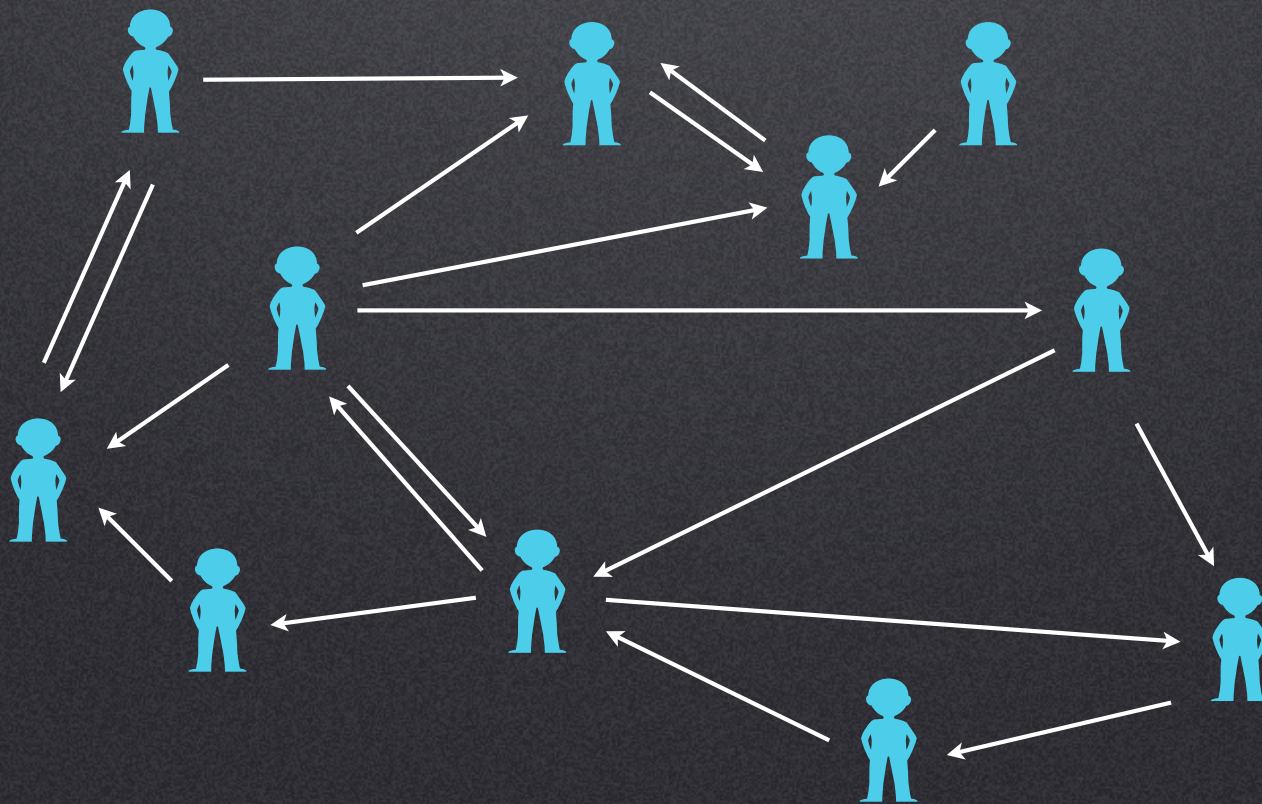


16 GB/day

Just tweet text!

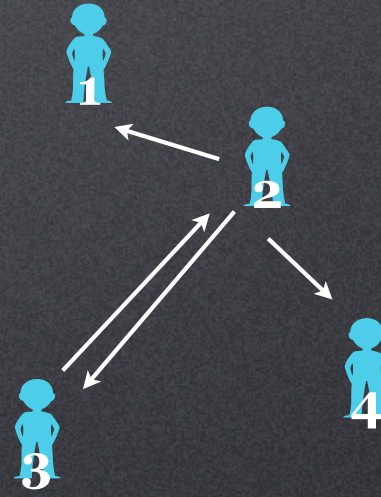


LOTS OF DELIVERIES



Digraph

Need to represent this



	1	2	3	4
1		✓		
2			✓	
3		✓		
4		✓		

Matrix

Naïve implementation is not scalable

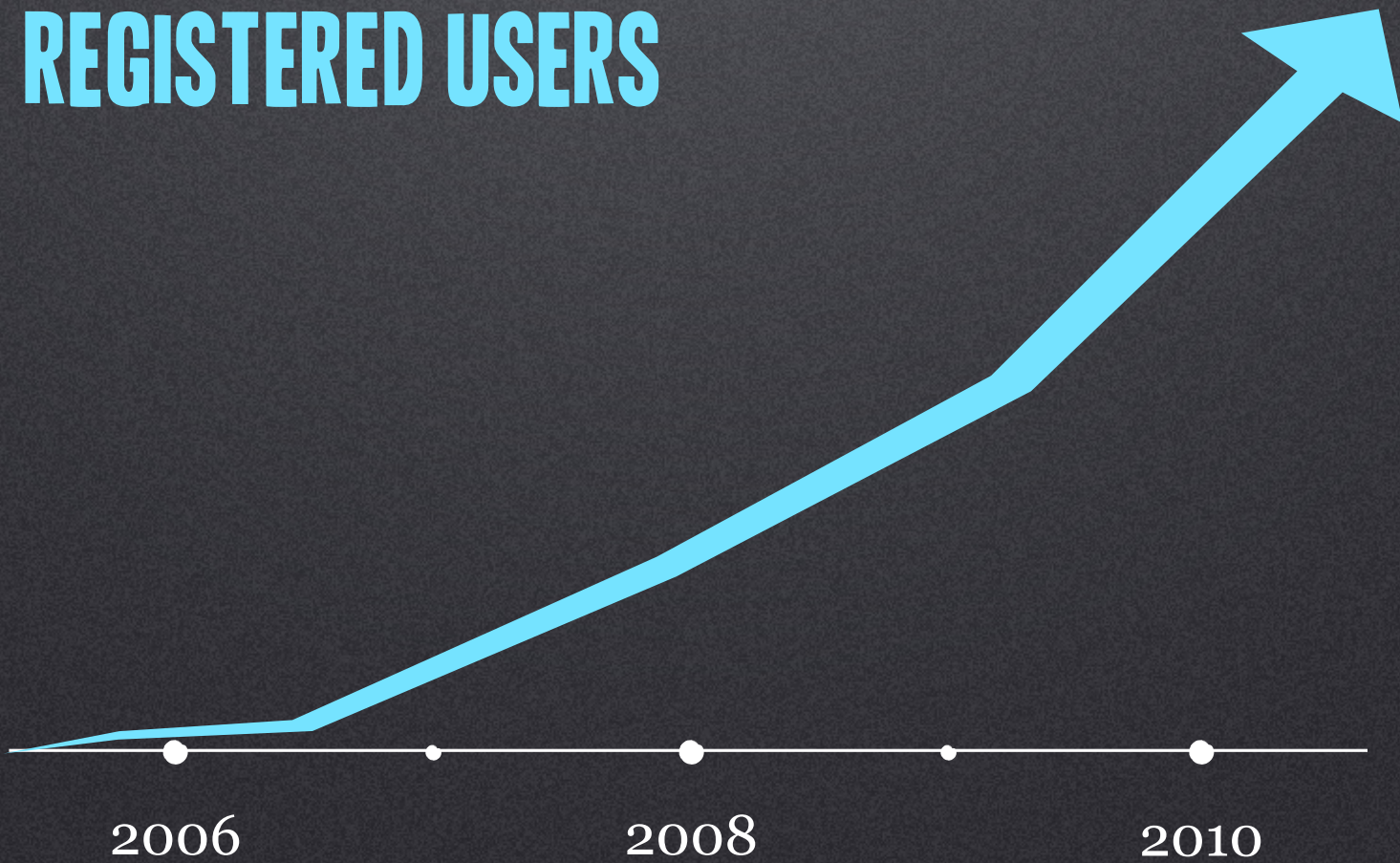


Photo used under Creative Commons from [jurvetson](#)

flockdb

Distributed graph database
High rate of CRUD operations
Complex set arithmetic queries
<http://github.com/twitter/flockdb>

160M REGISTERED USERS



USING THE SYSTEM

6B API calls
per day \approx 70,000 calls
per second

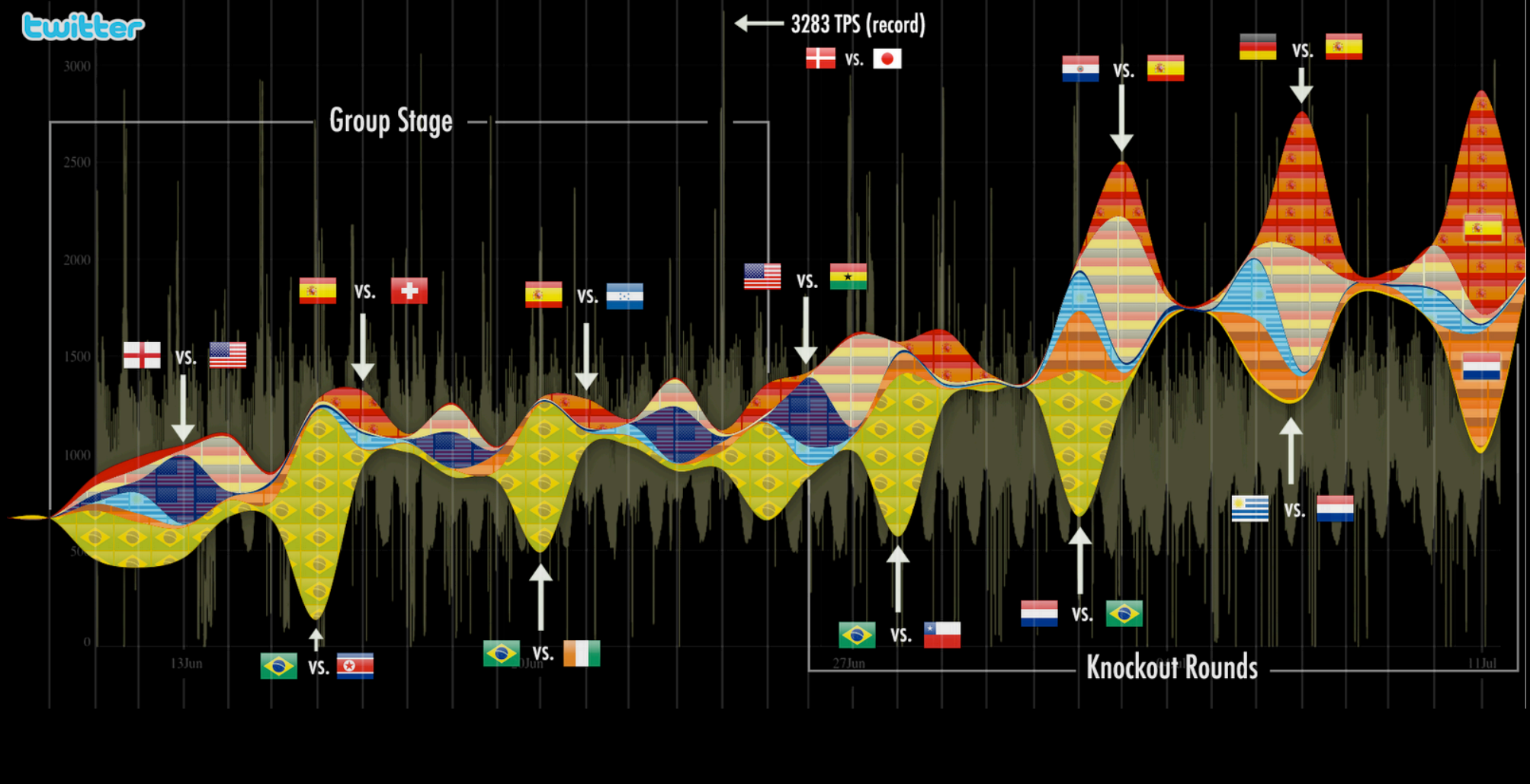
twitter

← 3283 TPS (record)

🇩🇰 vs. 🇯🇵

Group Stage

Knockout Rounds



WHERE DO WE WANT TO BE?

Today - 160M people generate ~1000 TPS

Tomorrow - we want to support half the world and all its devices

(right now, there are 6B people and 5B phones)

QUESTIONS?

Follow me at
twitter.com/raffi

