

Gossip and Inconsistency in a Search Engine

Knut Magne Risvik

Distributed Index Serving 101

Document sharded



All machines engage.
Completes the search.
Merge results.

Term sharded



A few machines engage.
Cross talk.
Works on disks...

Search Funnel

All documents



100B+

Matches



1000+
Boolean* search
Prune losers

Ranked results



10s
Complex ranker
Find winners
Error correcting

Updates



Organized
Local to each bucket



Slightly more complex
Global

Lame dictator way



Joe ·

I believe in the right to keep and bear arms. One Facebooker had challenged all gun owners to put this on their wall. The Supreme Court of the United States has affirmed that the right to keep and bear arms is a fundamental right of the individual. **If you** believe in the 2nd Amendment, and **you are** not afraid to show it. Re-post.

6 hours ago · Like · Comment

3 people like this.

Write a comment...



Jennifer ·

If you are doing a review for me please fill this out whenever **you do** your review so I can keep up please! <http://tinyurl.com/4c4zoz2> Also I only have 2 people showing no shipment yet and those 2 ladies already know. **If you** don't get your product by Saturday please let me know. Thank!

Amazon Review Status
tinyurl.com

7 hours ago · Like · Comment · Share



Carolyn ·

Please put this on your status **if you** know or love someone who is gay. My wish for 2011 is that people will understand that being gay is not a DISEASE nor a CHOICE - people who **are gay** are not looking for a CURE but ACCEPTANCE and EQUAL RIGHTS!! 93% won't copy and paste this. Will **you** make it your status for at least one hour? Promote LOVE And Equality for all!

7 hours ago · Like · Comment

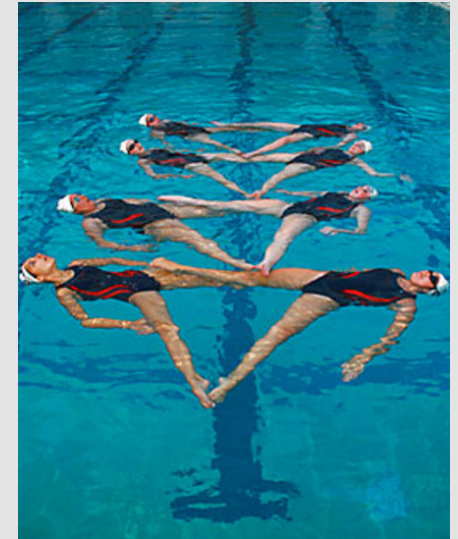
Beth Thomas likes this.

Write a comment...

Get updates



Drain traffic



Synchronize

Gossip way

Update rank data (easier)



Partition updates
Spread mutations



Immediately serve



Gossip about your
state and replicas

Yummies and Yucks

No need for the dictator

Faster time to serve

Failure resilient (dictators tend to be unreliable)

False positives

Always inconsistent index

Error correction needed, might lose some

Still need some sync (rank data first)

Final thought?

- What is the real In-consistency of the data for a search engine?
- Still need to measure to possible effect from false negatives.