Side by Side with Elasticsearch & Solr

Part 2 - Performance & Scalability

Radu Gheorghe
Rafał Kuć

sematext

Who are we?

Radu

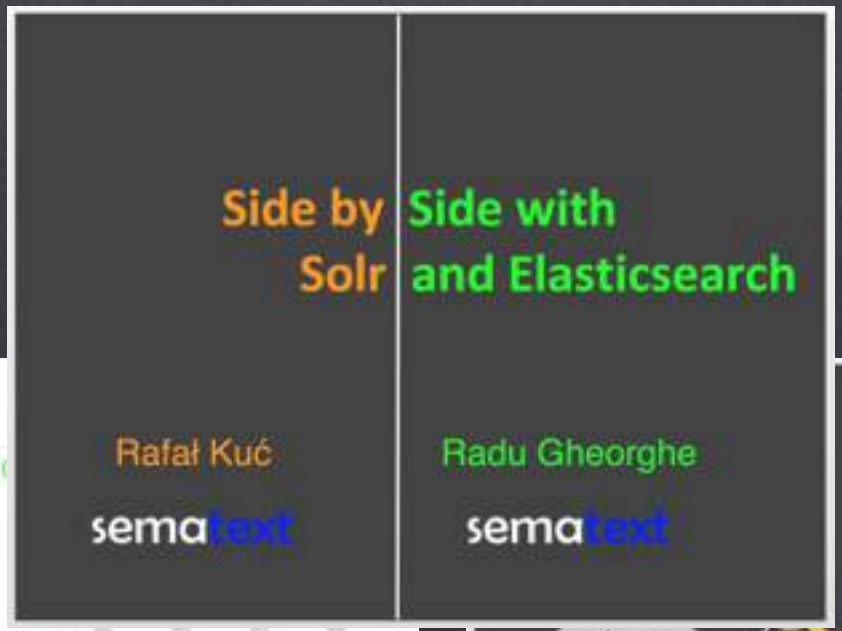
cftweia

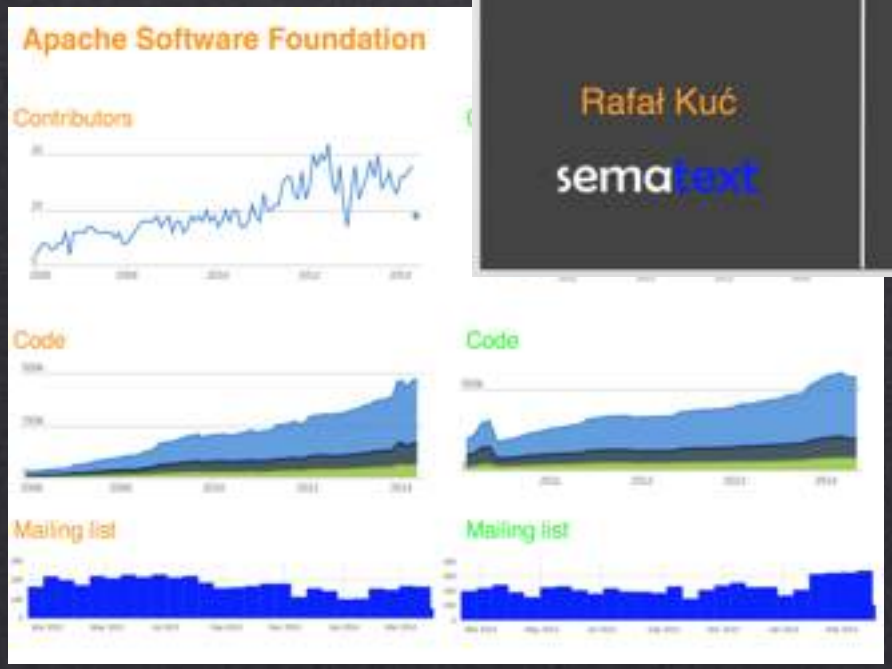
sematext

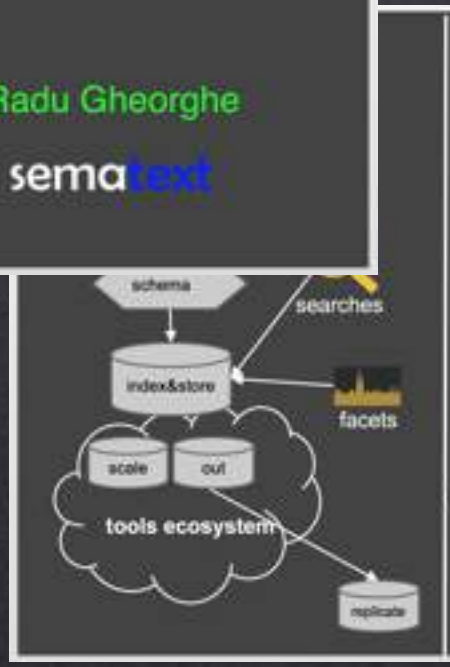
Logsene

Rafał

Last year

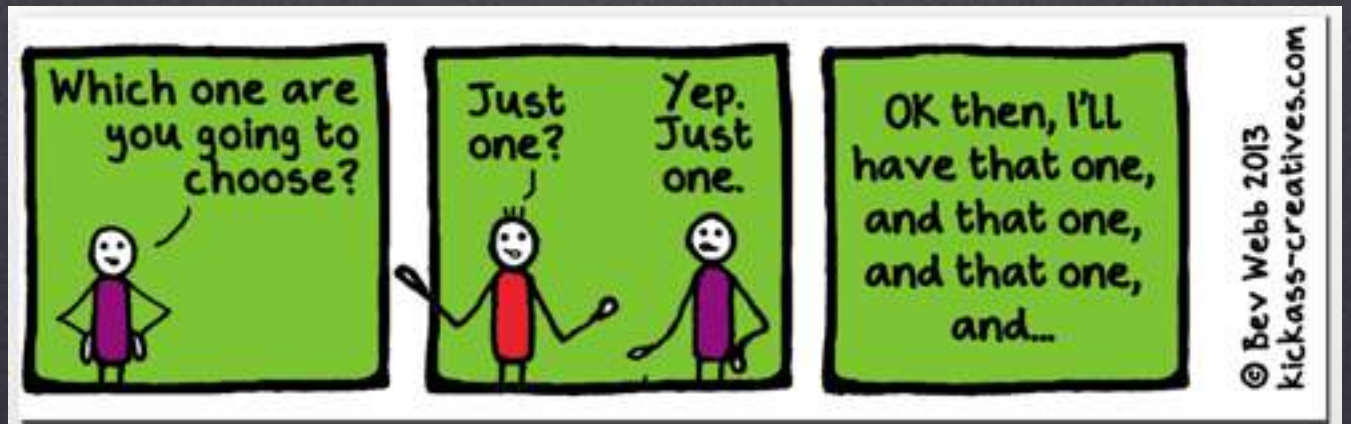

Side by Side with
Solr and Elasticsearch
Rafał Kuć
sematext
Radu Gheorghe
sematext


Apache Software Foundation
Contributors
Code
Code
Mailing list
Mailing list


schema
searches
index&store
facets
scale
out
tools ecosystem
replicate


Agenda
documents
queries
mapping
searches
index&store
aggregations
scale
out
tools ecosystem
percolations
backup

Last year conclusions



most projects work well with either

many small differences, few show-stoppers

choose the best. for your use-case.

One year progress

Facet by function
https://issues.apache.org/jira/browse/SOLR-1581

Analytics component
https://issues.apache.org/jira/browse/SOLR-5302

Solr as standalone application
https://issues.apache.org/jira/browse/SOLR-4792

top_hits aggregation
https://github.com/elastic/elasticsearch/pull/6124

minimum_should_match on has_child
https://github.com/elastic/elasticsearch/pull/6019

filters aggregation
https://github.com/elastic/elasticsearch/pull/6118

That's not all

JSON facets
https://issues.apache.org/jira/browse/SOLR-7214

Backup + restore
https://issues.apache.org/jira/browse/SOLR-5750

Streaming aggregations
https://issues.apache.org/jira/browse/SOLR-7082

Cross data center replication
https://issues.apache.org/jira/browse/SOLR-6273

Moving averages aggregation
https://github.com/elastic/elasticsearch/pull/10002

Computation on aggregations
https://github.com/elastic/elasticsearch/pull/9876

Cluster state diff support
https://github.com/elastic/elasticsearch/pull/6295

Shadow replicas
https://github.com/elastic/elasticsearch/pull/8976

sematext

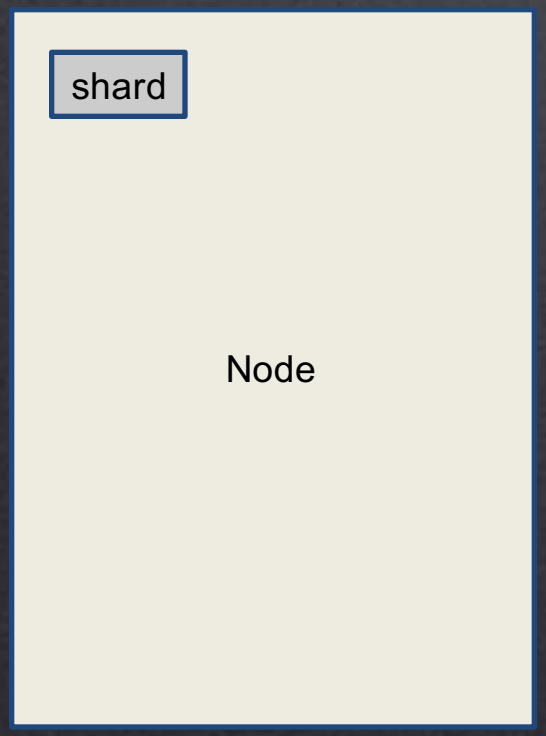
This year's agenda

Horizontal scaling

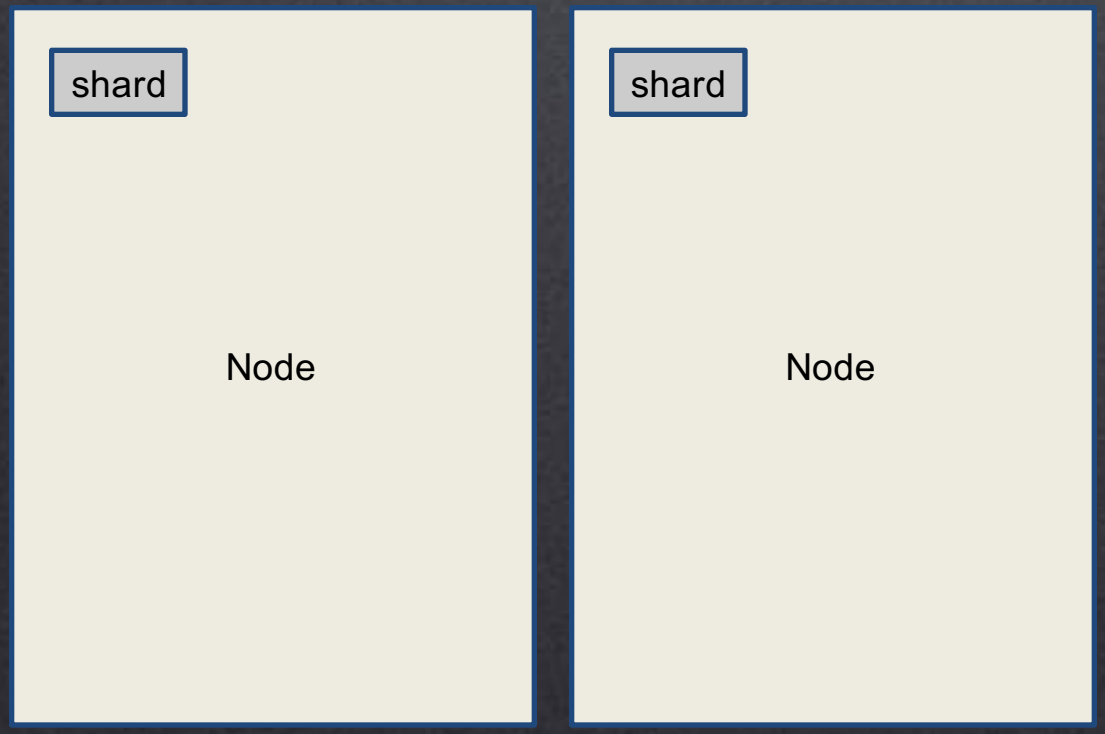
Products use-case

Logs use-case

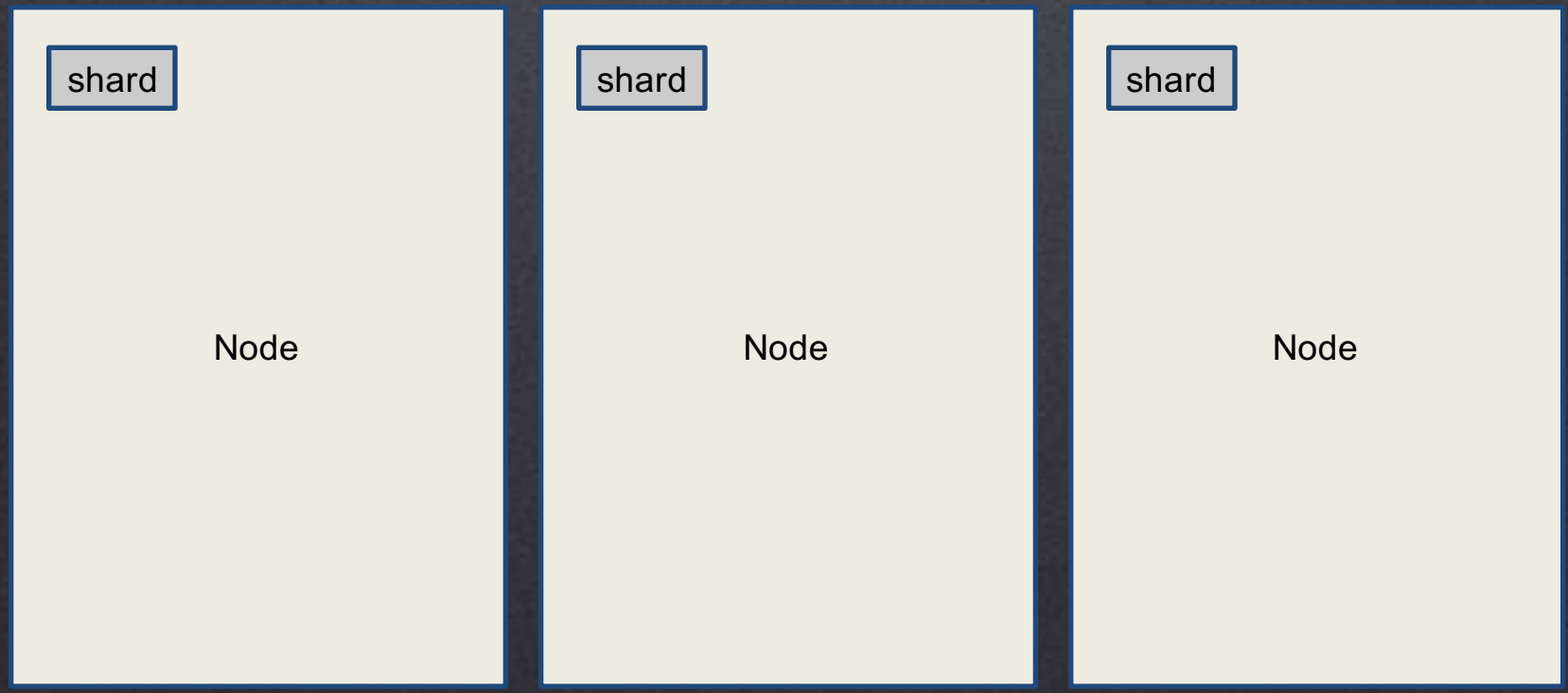
Horizontal scaling in theory



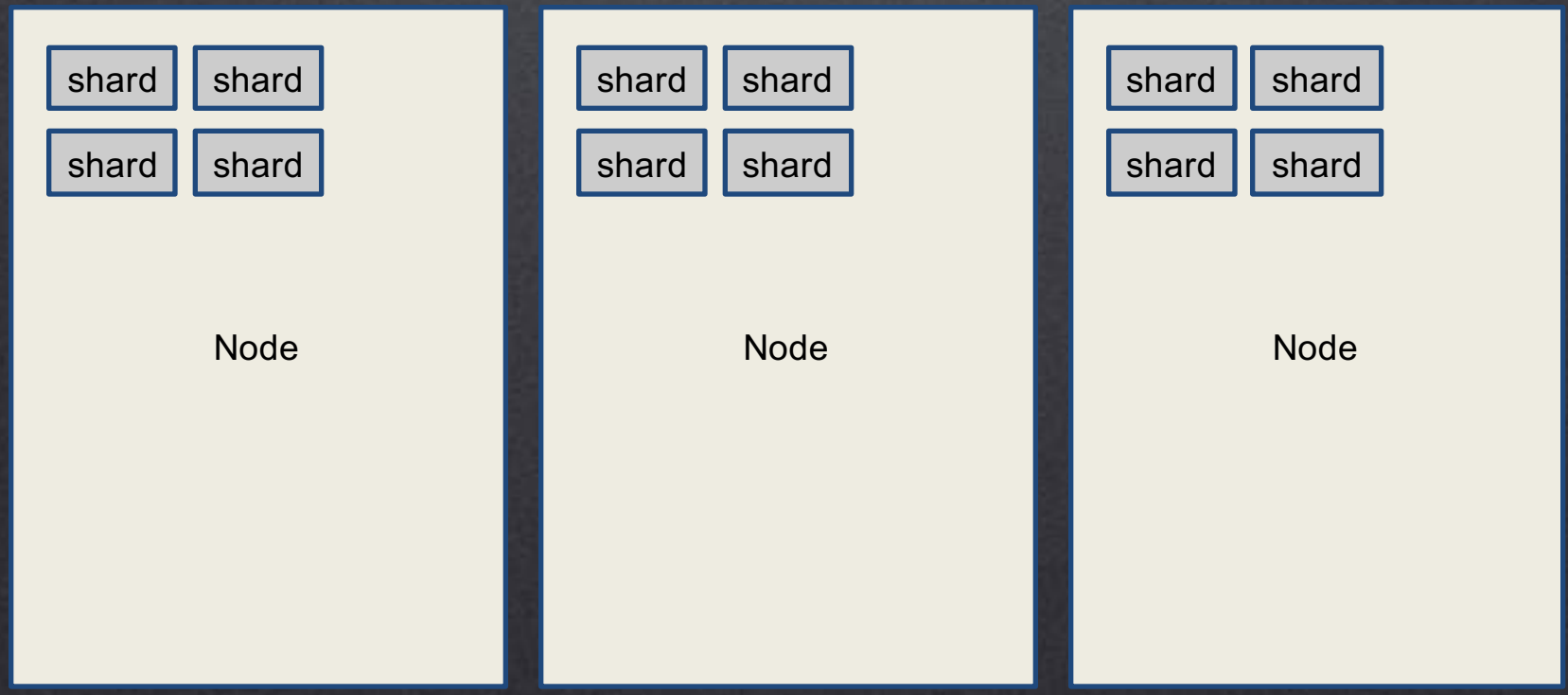
Horizontal scaling in theory



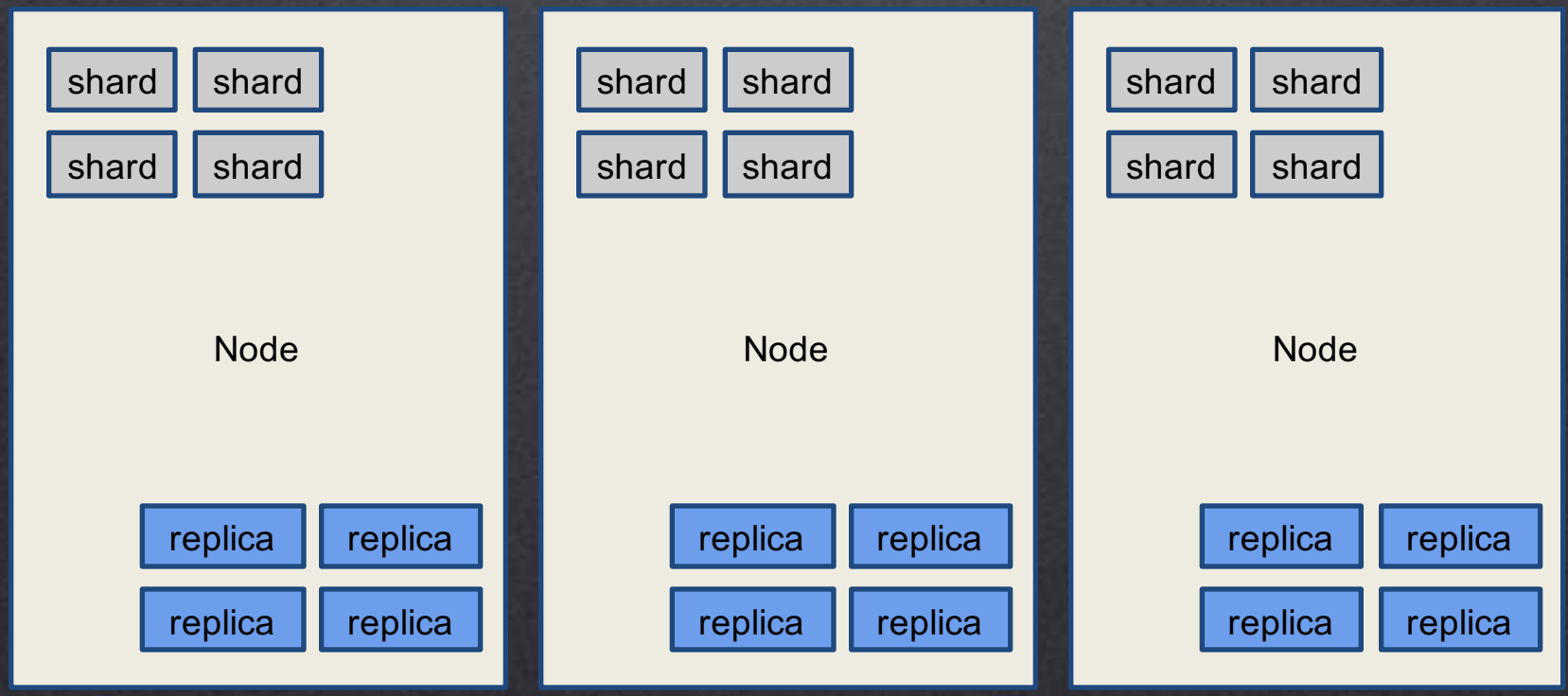
Horizontal scaling in theory



Horizontal scaling in theory



Horizontal scaling in theory



Horizontal scaling - the API

Create / remove replicas on the fly - Collections API

Moving shards around the cluster using add / delete replica

Shard splitting using Collections API

Migrating data with a given routing key to another collection using API

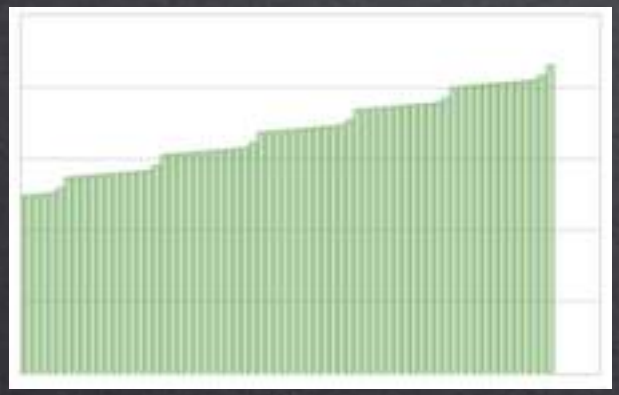
Create / remove replicas on the fly - Update Indices Settings

Moving shards around the cluster using Cluster Reroute API

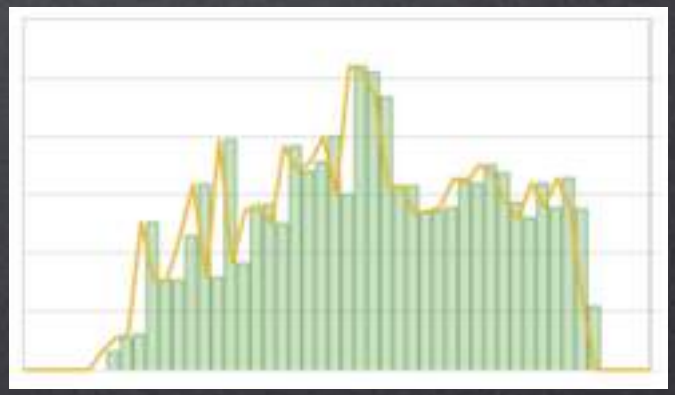
Automatic shard balancing by default

Shard allocation awareness & rule based shard placement

The products - assumptions



Steady data growth



Spikes in traffic

Large QPS

Common, known data structure

Hardware and data

vs

2 x EC2 c3.large instances
(2vCPU, 3.5GB RAM,
2x16GB SSD in RAID0)

Wikipedia

Test requests

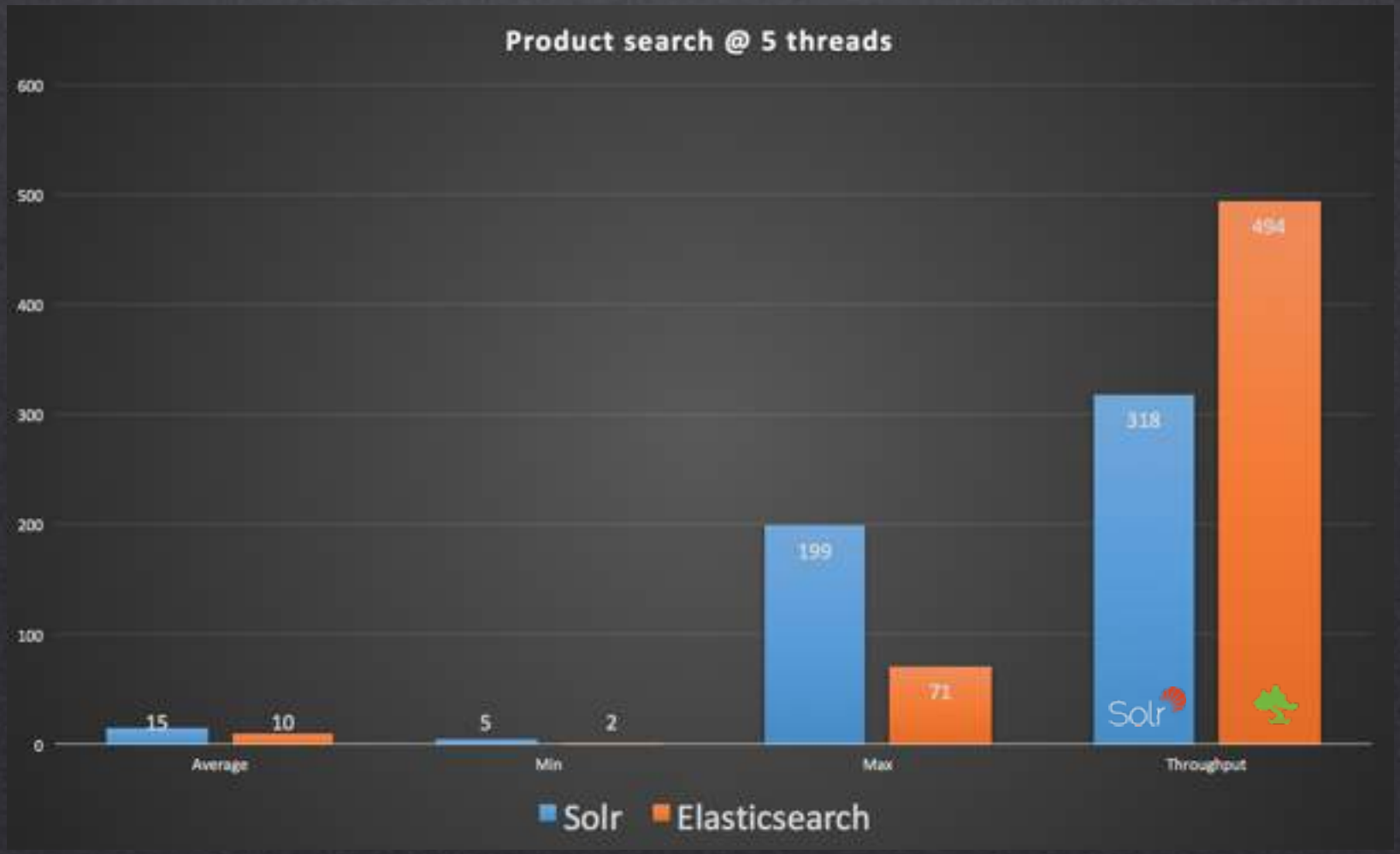
One, common query

Dictionary of common and uncommon terms

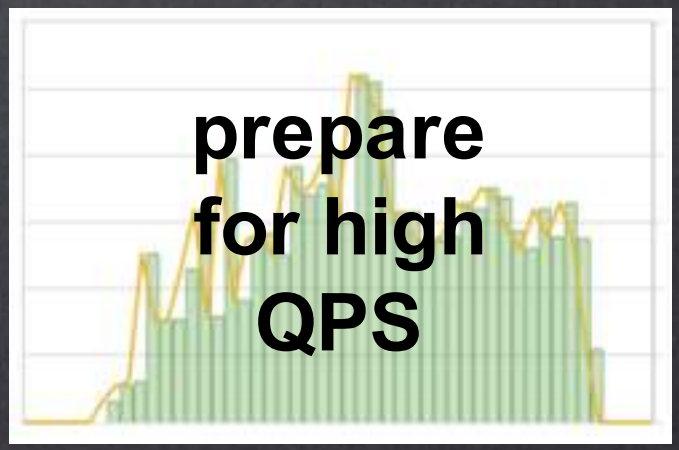
JMeter hammering

https://github.com/sematext/berlin-buzzwords-samples/tree/master/2015

Product search @ 5 threads



The products - summary



Both are fast, but with wikipedia

Solr < 

The products - summary

plan for data growth

prepare for high QPS

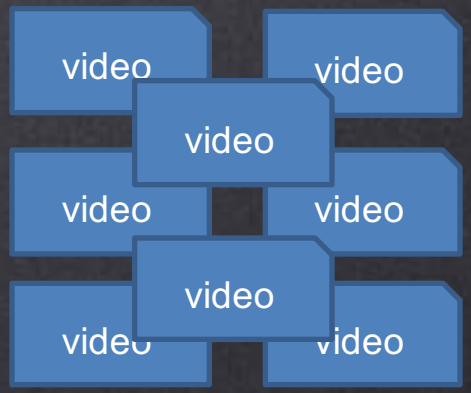
Both are fast, but with wikipedia

Solr < 

sematext

Hardware and data (2nd try)

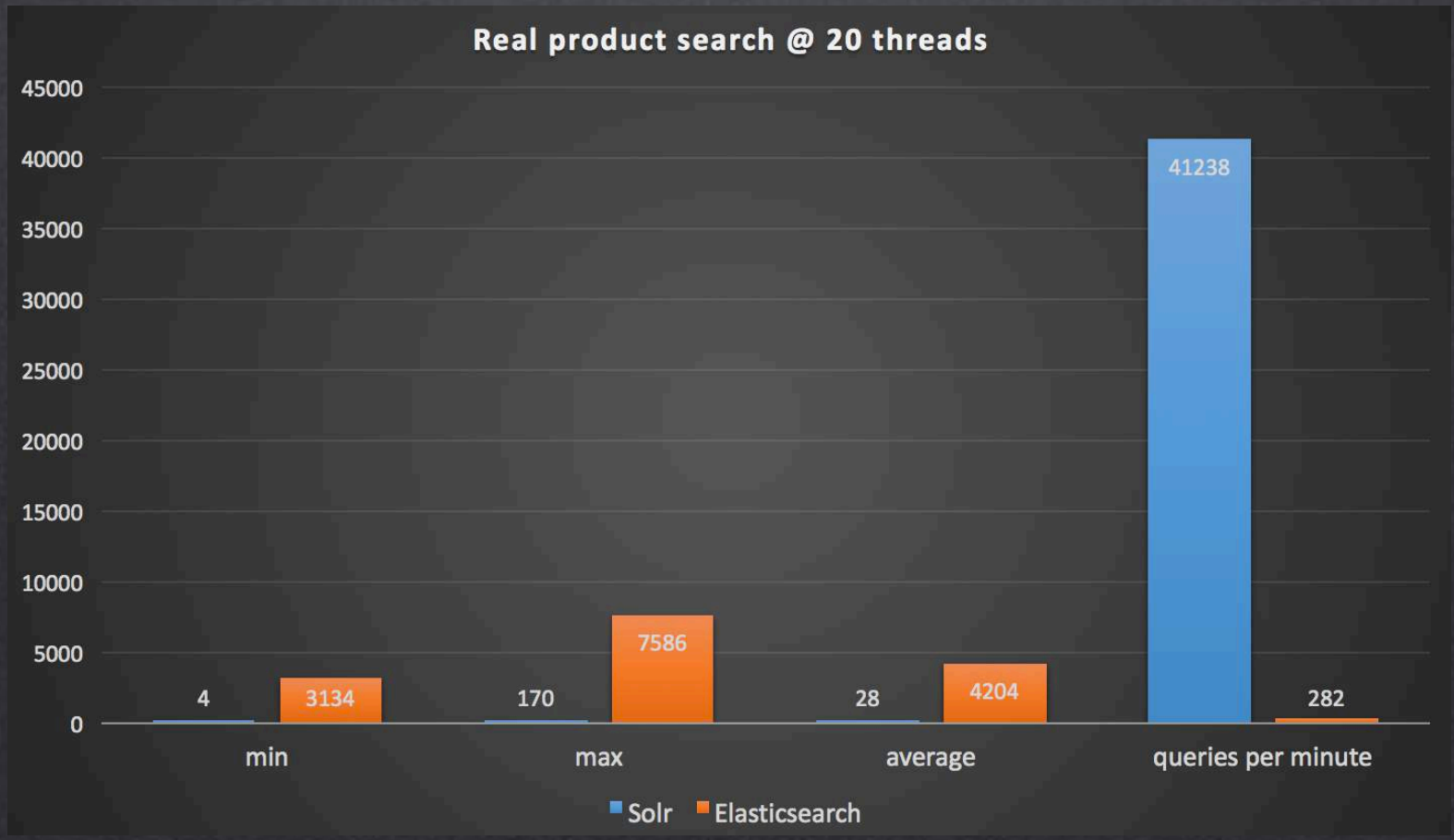
vs



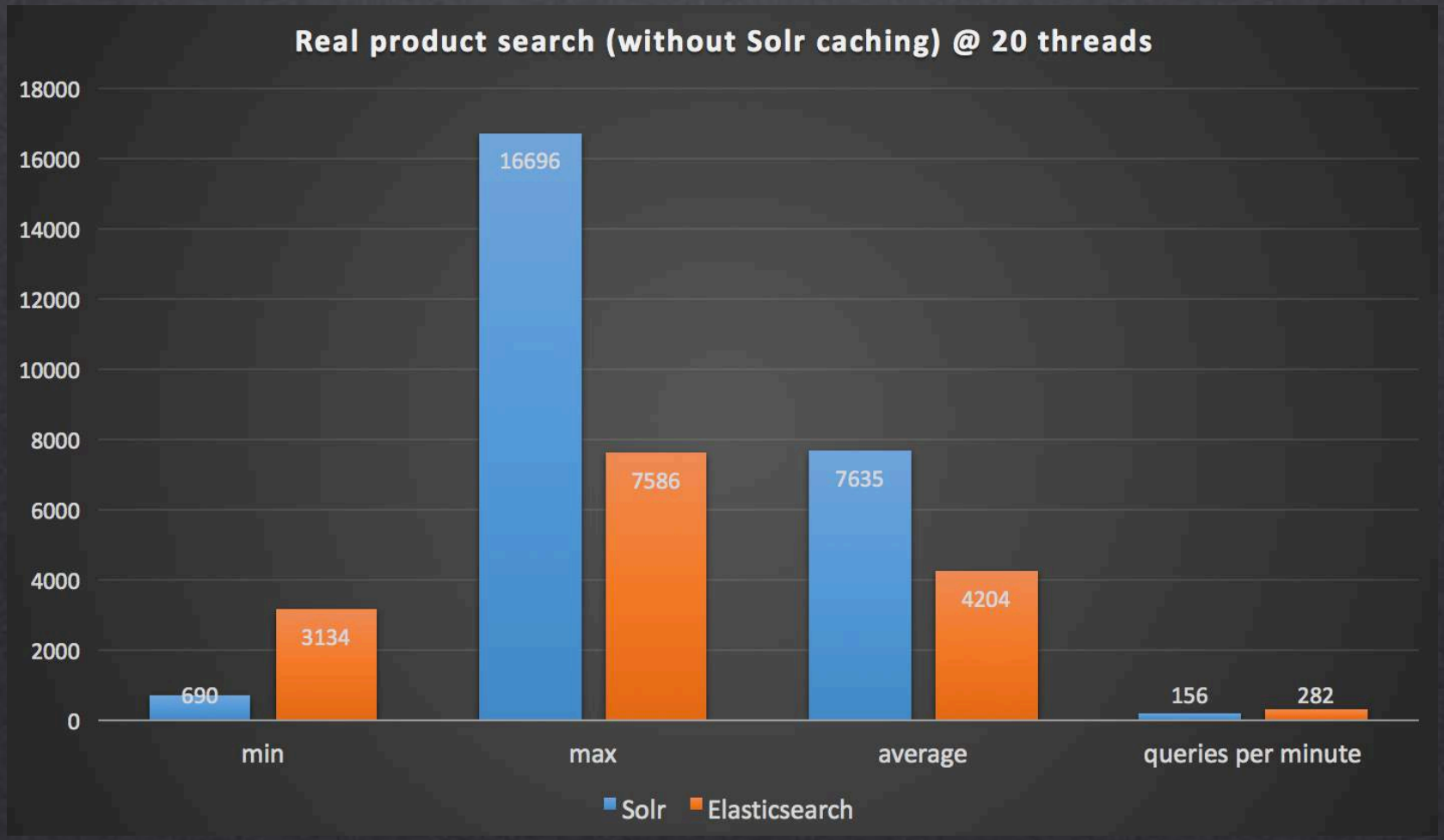
2 x EC2 c3.large instances
(2vCPU, 3.5GB RAM,
2x16GB SSD in RAID0)

Video
search

Real product search @ 20 threads



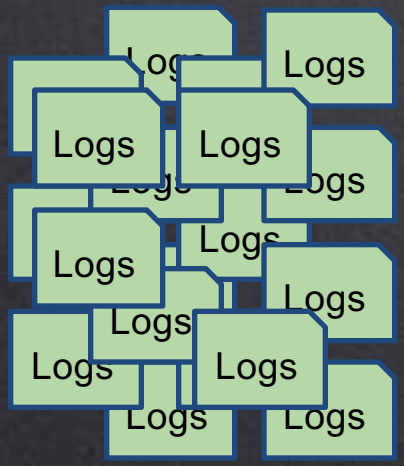
Real product search @ 20 threads



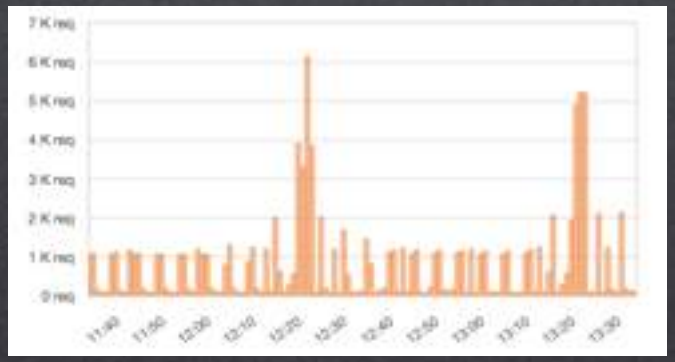
The products – real summary

With video both are fast and configuration matters

The logs - assumptions



Lots of data



Low query count

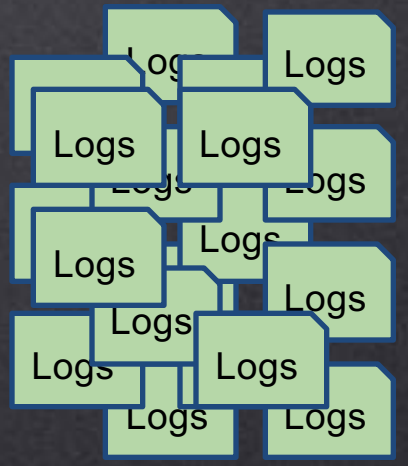
Time oriented data

No updates

sematext

Hardware and data

vs



2 x EC2 c3.large instances
(2vCPU, 3.5GB RAM,
2x16GB SSD in RAID0)

Apache logs

Tuning

soft autocommit: 5s

hard autocommit: 200mb

docValues: true

catch all field: on

refresh_interval: 5s

flush_threshold_size: 200mb

doc_values: true

_all: enabled

Test requests

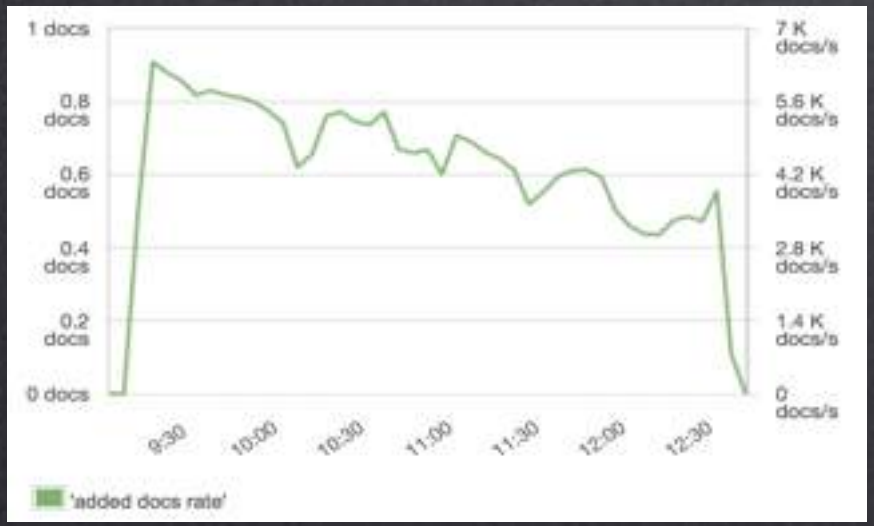
Filters	Aggregations/Facets
filter by client IP	date histogram 
filter by word in user agent	top 10 response codes 
wildcard filter on domain	# of unique IPs
	top IPs per response per time

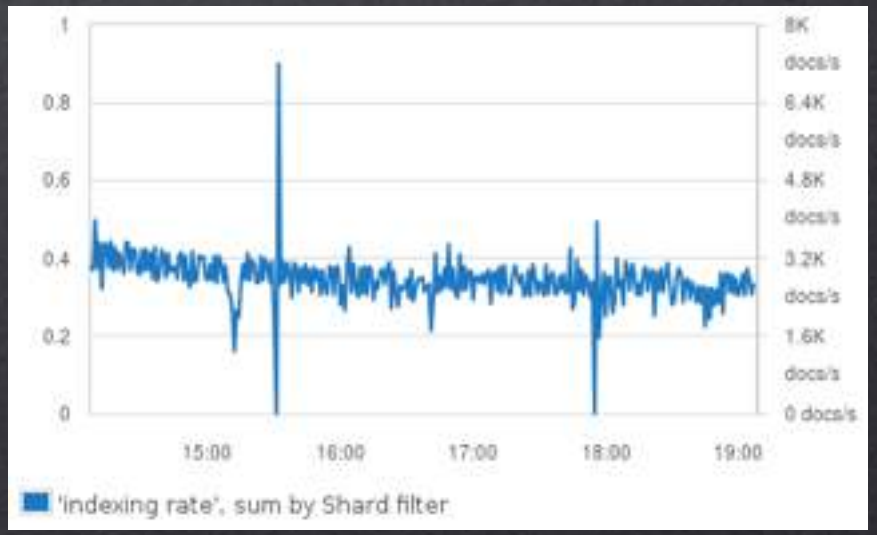
https://github.com/sematext/berlin-buzzwords-samples/tree/master/2015

Test runs

1. Write throughput
2. Capacity of a single index
3. Capacity with time-based indices on hot/cold setup

Single index write throughput

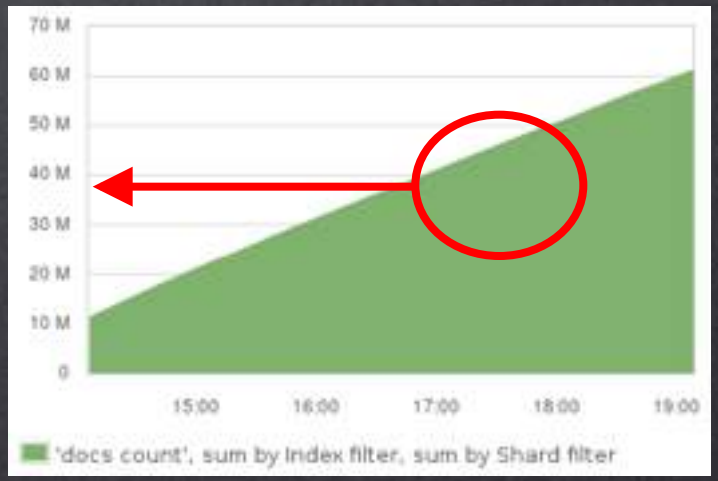


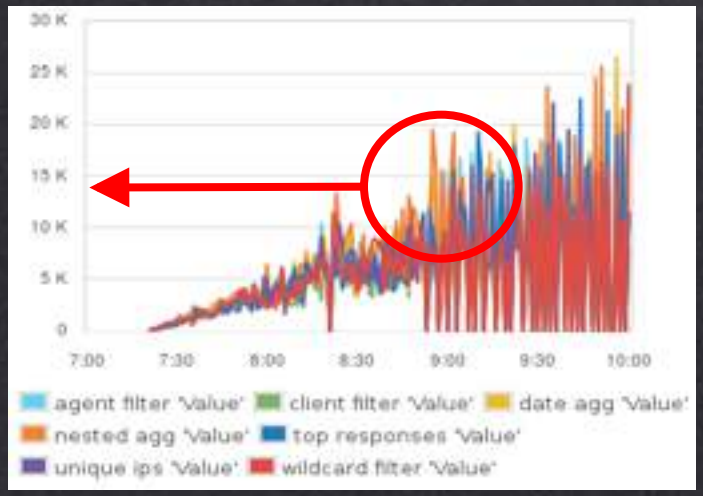


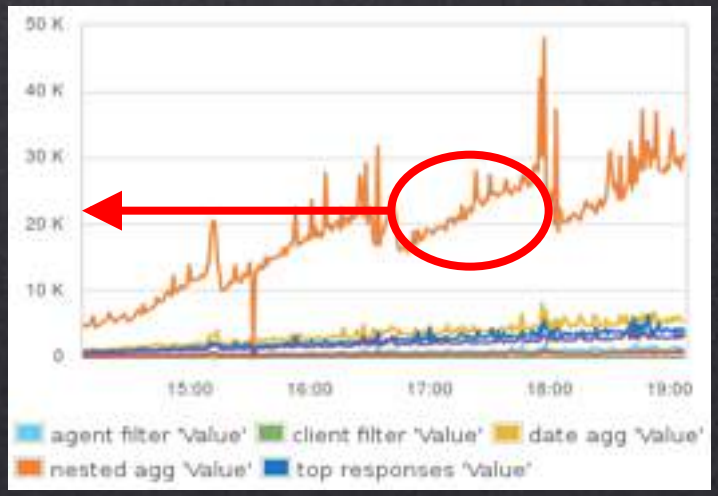
sematext

Solr

Single index @ max EPS

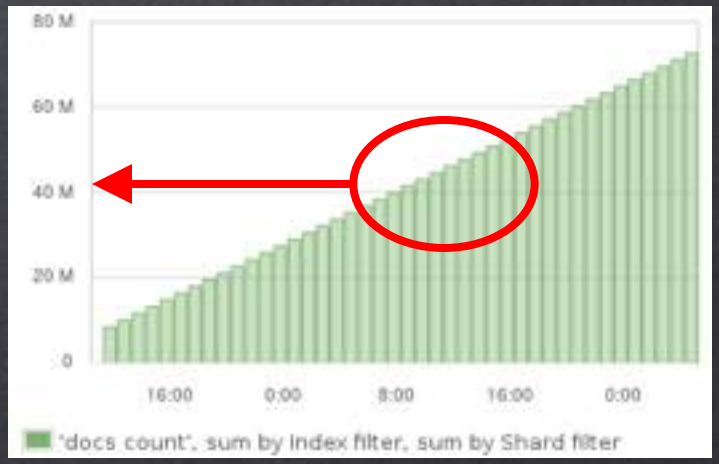


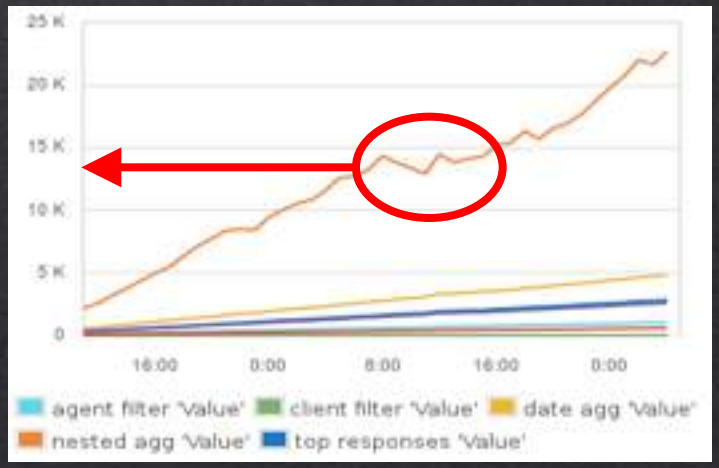




sematext Solr

Single index @ 400 EPS





sematext

Solr

Time-based indices

smaller indices
- lighter indexing
- easier to isolate hot data from cold data
- easier to relocate

bigger indices
- less RAM
- less management overhead
- smaller cluster state

Hot / Cold in practice

Solr

cron job does everything

creates indices in advance, on the hot nodes (createNodeSet property)

Optimizes old indices, creates N new replicas of each shard on the cold nodes

Removes all the replicas from the hot nodes

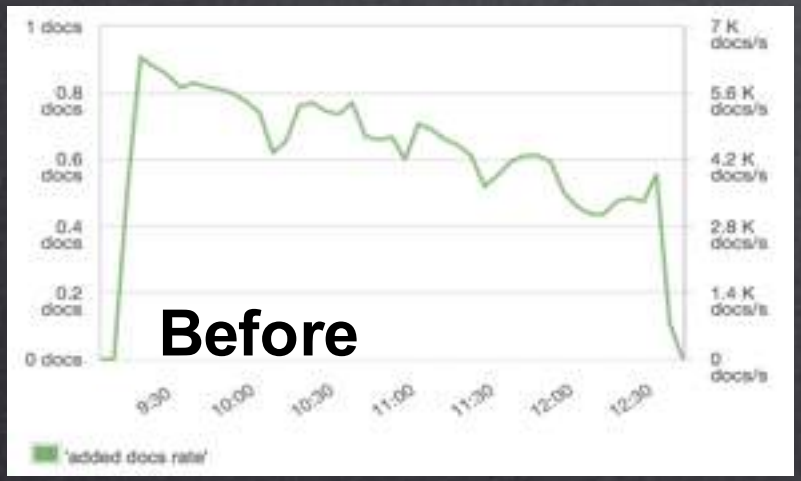
Uses node properties (e.g. node.tag)

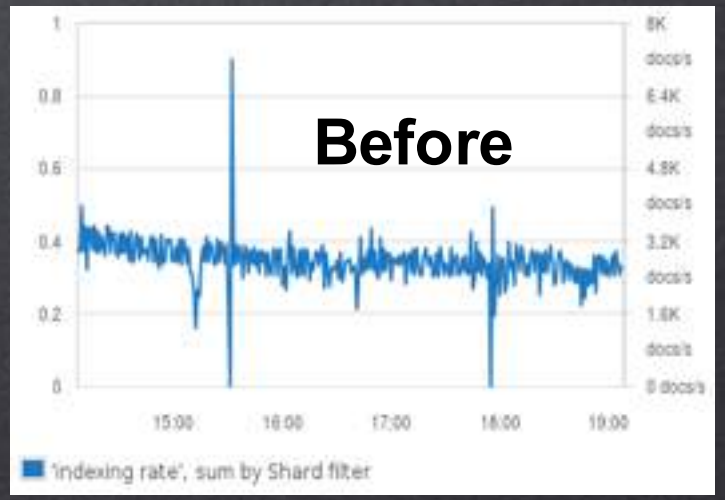
index templates + shard allocation awareness = new indices go to hot nodes automagically

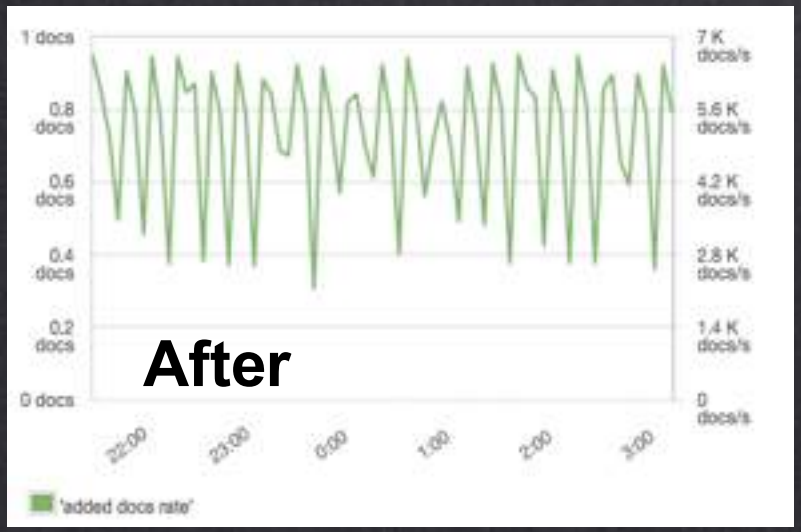
Cron job optimizes old indices and changes shard allocation attributes => shards get moved and Do on the cold nodes

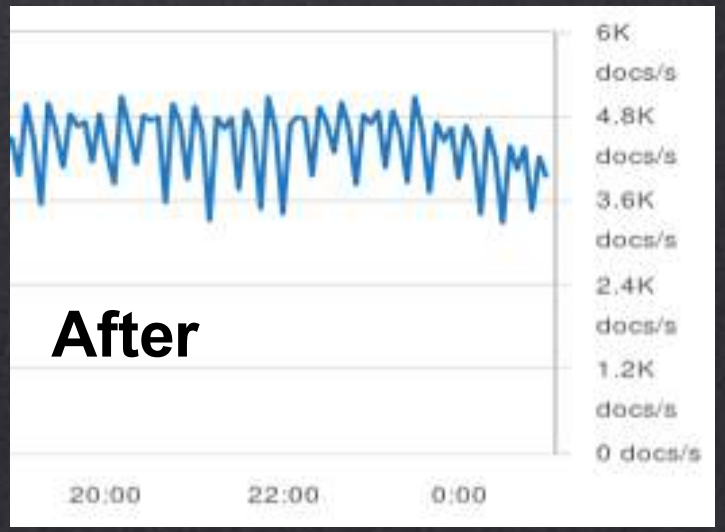
sematext

Time-based: 2 hot and 2 cold



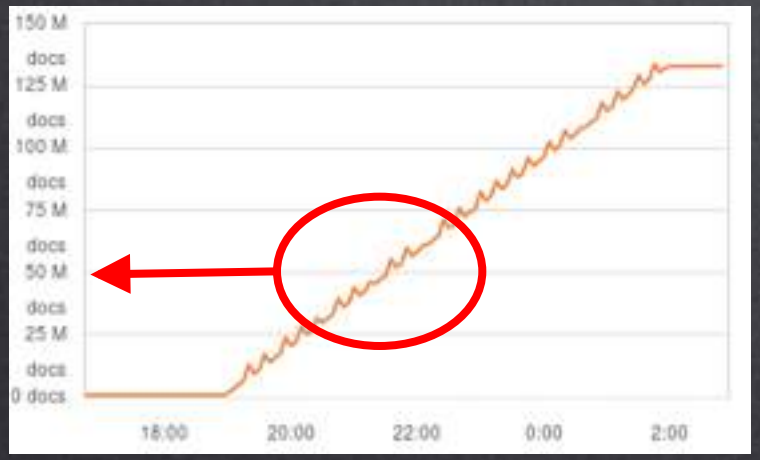


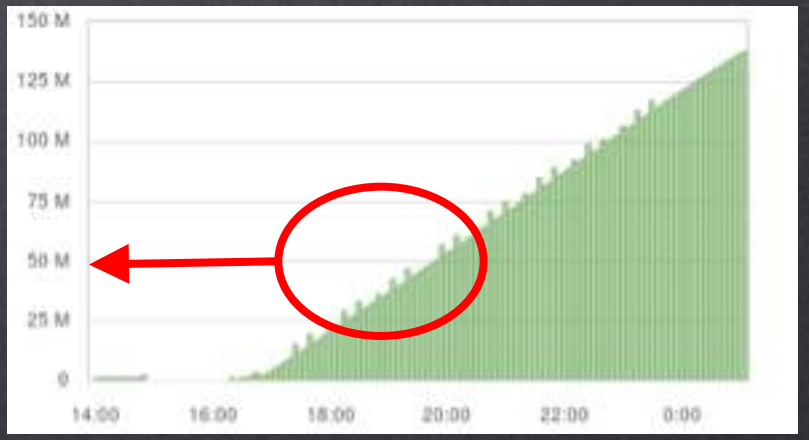


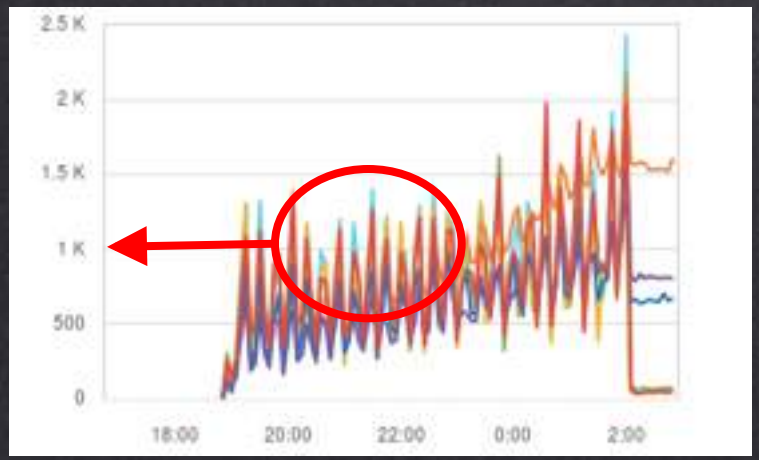


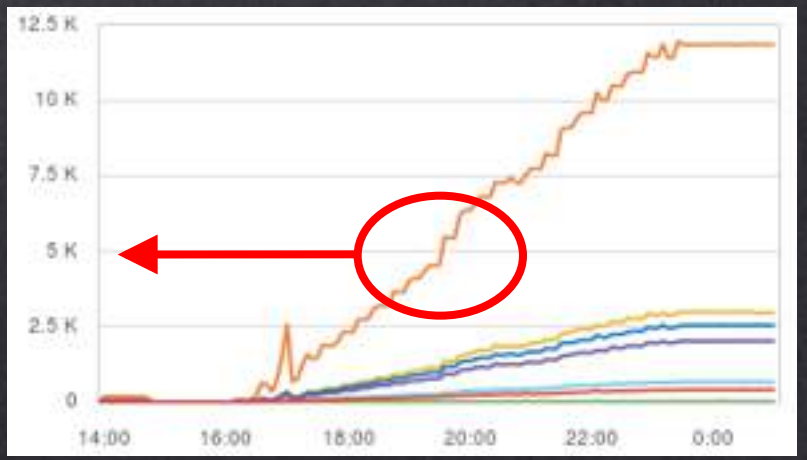
sematext Solr

Time-based: 2 hot and 2 cold









sematext Solr

The logs - summary

Solr

Faceting

Filtering

One summary to rule them all

Differences in configuration often matter more than differences in products

Before the tests we expected results to be totally opposite in both use cases

Do your own tests with your data and queries before jumping into conclusions

We are hiring

Dig Search?
Dig Analytics?
Dig Big Data?
Dig Performance?
Dig Logging?
Dig working with and in open – source?
We're hiring world – wide!

http://sematext.com/about/jobs.html

sematext

Call me, maybe?

Radu Gheorghe

@radu0gheorghe

radu.gheorghe@sematext.com

Rafał Kuć

@kucrafal

rafal.kuc@sematext.com

Sematext

@sematext

http://sematext.com

sematext