



# elasticsearch.

**Florian Hopf**  
**@fhopf**

# Agenda

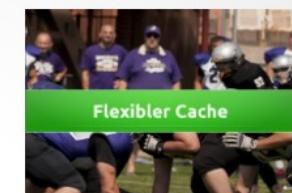
- Anwendungsfälle



Document Store



Volltextsuche



Flexibler Cache



Geo-Suche

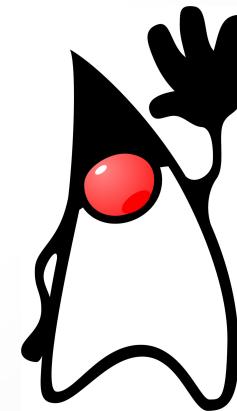


Log-Analyse



Analytics

- Integration in Java



# Agenda



Document Store



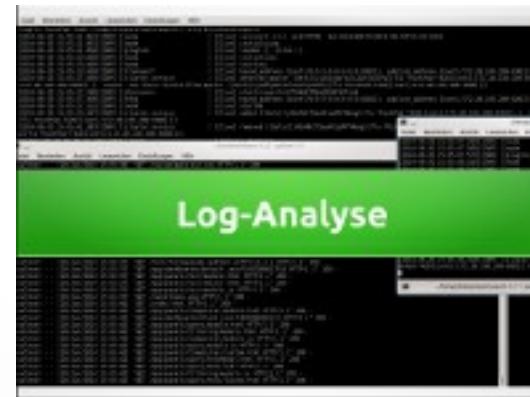
Volltextsuche



Flexibler Cache



Geo-Suche



Log-Analyse



Analytics

# Vorbereitung



# Installation

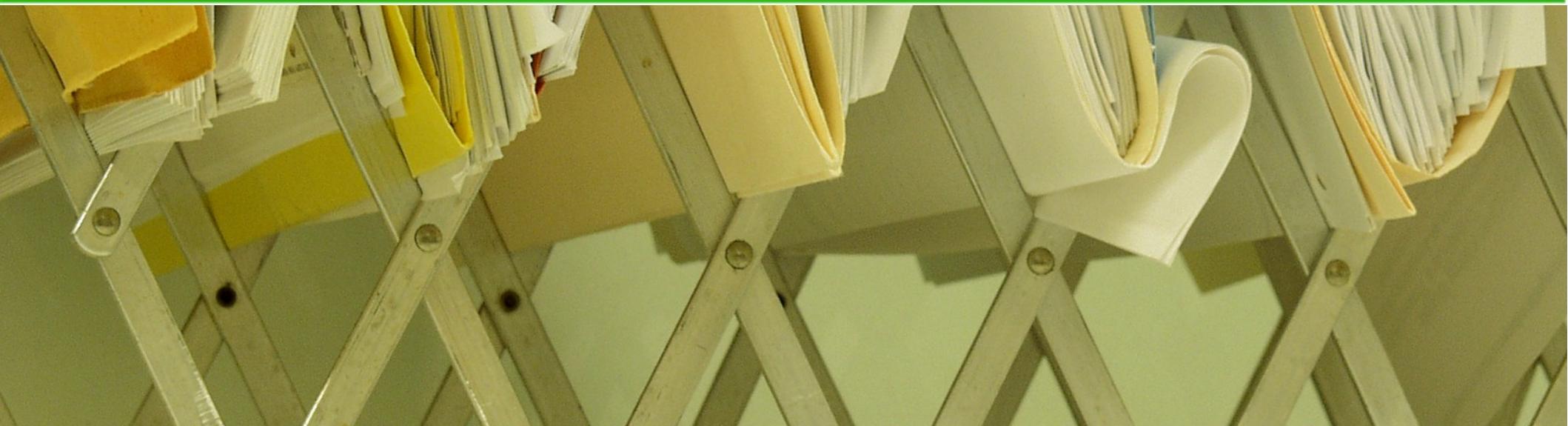
```
# download archive  
wget https://download.elastic.co/  
      elasticsearch/elasticsearch/elasticsearch-1.7.1.zip  
  
# zip is for windows and linux  
unzip elasticsearch-1.7.1.zip  
  
# on windows: elasticsearch.bat  
elasticsearch-1.7.1/bin/elasticsearch
```

# Zugriff

```
curl -XGET http://localhost:9200
{
  "status" : 200,
  "name" : "Bird-Man",
  "cluster_name" : "elasticsearch",
  "version" : {
    "number" : "1.7.1",
    "build_hash" : "b88f43fc40b0bcd7f173a1f9ee2e97816",
    "build_timestamp" : "2015-07-29T09:54:16Z",
    "build_snapshot" : false,
    "lucene_version" : "4.10.4"
  },
  "tagline" : "You Know, for Search"
}
```



# Document Store



# Document

```
{  
  "title" : "Anwendungsfälle für Elasticsearch",  
  "speaker" : "Florian Hopf",  
  "date" : "2015-09-15T19:00:00.000Z",  
  "tags" : ["Java", "Lucene"],  
  "conference" : {  
    "name" : "JUG-KL",  
    "city" : "Kaiserslautern"  
  }  
}
```

# Speichern

```
curl -XPOST http://localhost:9200/conferences/talk/  
--data-binary @talk-example.json
```

```
{  
  "_index": "conferences",  
  "_type": "talk",  
  "_id": "GqjY7l8sTxa3jLaFx67_aw",  
  "_version": 1,  
  "created": true  
}
```

# Speichern

```
curl -XPOST http://localhost:9200/conferences/talk/  
--data-binary @talk-example.json
```

```
{  
  "_index":"conferences",  
  "_type":"talk",  
  "_id":"GqjY718sTxa3jLaFx67_aw",  
  "_version":1,  
  "created":true  
}
```



Index

# Speichern

```
curl -XPOST http://localhost:9200/conferences/talk/  
--data-binary @talk-example.json
```

```
{  
  "_index": "conferences",  
  "_type": "talk",  
  "_id": "GqjY718sTxa3jLaFx67_aw",  
  "_version": 1,  
  "created": true  
}
```

↑  
Type

# Lesen

```
curl -XGET http://localhost:9200/conferences/talk/  
      GqjY7l8sTxa3jLaFx67_aw?pretty=true  
{  
  "_index" : "conferences",  
  [...]  
  "_source":{  
    "title" : "Anwendungsfälle für Elasticsearch",  
    "speaker" : "Florian Hopf",  
    "date" : "2015-09-15T19:00:00.000Z",  
    "tags" : ["Java", "Lucene"],  
    "conference" : {  
      "name" : "JUG-KL",  
      "city" : "Kaiserslautern"  
    }  
  }  
}
```

# Sharding

- Aufteilen eines Index in mehrere Teile
  - Default: 5 Shards pro Elasticsearch-Index
- Mehrere ES-Instanzen bilden einen Cluster
  - Automatische Verteilung auf die Knoten

# Sharding

13:07:35 - elasticsearch

   	conferences shards: 5 * 2   docs: 1   size: 4.18KB
<p>★ Anole</p> <p>flo-ThinkPad-T420 - inet[/192.168.2.143:9300]</p> <p>docs: 0 size: 369.00b</p> <p>heap disk cpu</p>	<div style="display: flex; justify-content: space-around;"><div>1</div><div>3</div><div>4</div></div>
<p>★ Alyosha Kravinoff</p> <p>flo-ThinkPad-T420 - inet[/192.168.2.143:9302]</p> <p>docs: 1 size: 3.90KB</p> <p>heap disk cpu</p>	<div style="display: flex; justify-content: space-around;"><div>0</div><div>1</div><div>2</div></div>
<p>★ Talia Josephine Wagner</p> <p>flo-ThinkPad-T420 - inet[/192.168.2.143:9301]</p> <p>docs: 1 size: 3.99KB</p> <p>heap disk cpu</p>	<div style="display: flex; justify-content: space-around;"><div>0</div><div>2</div><div>3</div><div>4</div></div>
<p>! unassigned shards</p>	

# Recap

- Einfache Speicherung von JSON-Dokumenten
- Index und Type
- Sharding für große Datenmengen
- Verteilung ist First-Class-Citizen

# Nutzer

- Atlassian HipChat
  - <http://highscalability.com/blog/2014/1/6/how-hipchat>
- OpenCrowd
  - <https://www.elastic.co/use-cases/open-crowd>

# Volltextsuche



← → C <https://www.google.de/#q=elasticsearch>

Google

Microphone icon Search icon

- elasticsearch
- elasticsearch
- elasticsearch head
- elasticsearch tutorial
- elasticsearch delete index

[Weitere Informationen](#)

Cookies helfen uns bei der Bereitstellung unserer Dienste. Durch die Nutzung unserer Dienste erklären Sie sich damit einverstanden, dass wir Cookies setzen.

[Mehr erfahren](#) OK

## Elasticsearch.org Open Source Distributed Real Time ...

[www.elasticsearch.org/](http://www.elasticsearch.org/) ▾ Diese Seite übersetzen

Elasticsearch is a powerful open source search and analytics engine that makes data easy to explore.

[Download](#) - [Overview](#) - [Resources](#) - [Guide](#)

## Volltextsuche mit Elasticsearch | heise Developer

[www.heise.de/developer/.../Volltextsuche-mit-ElasticSearch-1920454.htm](http://www.heise.de/developer/.../Volltextsuche-mit-ElasticSearch-1920454.htm) ▾

26.07.2013 - Wenn es um die Volltextsuche in großen Datenmengen geht, führte lange kein Weg an Apache Solr vorbei. Seit 2010 steht mit Elasticsearch ...

[Abfragen und Schemafreiheit](#) - [Suchen und erweitern](#) - [Fazit](#)

## Elasticsearch.com Real Time Data Search and Analytics ...

[www.elasticsearch.com/](http://www.elasticsearch.com/) ▾ Diese Seite übersetzen

Elasticsearch Marvel picks up every operating metric Elasticsearch exposes through its APIs and visualizes it instantly. It's free for development use and a ...

## Elasticsearch - Wikipedia, the free encyclopedia

[en.wikipedia.org/wiki/Elasticsearch](http://en.wikipedia.org/wiki/Elasticsearch) ▾ Diese Seite übersetzen

Elasticsearch is a search server based on Lucene. It provides a distributed, multitenant-capable full-text search engine with a RESTful web interface and ...  
[Lucene](#) - [Multitenancy](#) - [Compass Project](#) - [List of information retrieval](#) ...



Alle Kategorien ▾

Suche

Alle Kategorien ▾

elasticsearch |

Suchvorschläge

Fremdsprachige Bücher

Erweiterte Suche

elasticsearch server in Alle Kategorien

7 Ergebnisse für Fremdsprachige Bücher : Comp

elasticsearch server in Englische Bücher

elasticsearch server in Fremdsprachige Bücher Trade-In

Ergebnisse anzeigen  
für< Alle Kategorien  
< Fremdsprachige Bücher  
< Computer & Internet  
< Programmieren  
**Sprachen & Tools**

Filtern nach

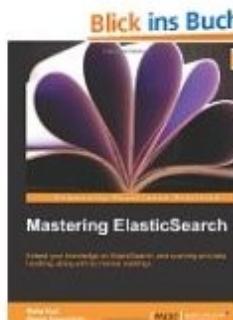
**Versandoption** (Was ist das?)  
Kostenlose Lieferung möglich**Sprachen von  
fremdsprachigen Büchern**  
 Englisch (7)**Neuheiten**Letzte 30 Tage  
Letzte 90 Tage  
Demnächst (1)**Durchschn. Kundenrezension**  
 & mehr (1)  
 & mehr (2)  
 & mehr (2)  
 & mehr (2)**Preis**

10 - 20 EUR (1)

20 - 50 EUR (6)



Andere Angebote - Taschenbuch

**EUR 43,00** neu (8 Angebote)**EUR 102,92** gebraucht (3 Angebote)**Mastering ElasticSearch** von Rafal Kuc und Marek Rogozinski von Packt Publishing (25. Oktober 2014)

Kostenlose Lieferung möglich.

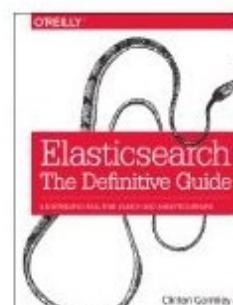
**EUR 46,00** Taschenbuch 

Lieferung bis zum Samstag, 21. Juni

**EUR 18,74** Kindle Edition

Jetzt als Download verfügbar.

Andere Angebote - Taschenbuch

**EUR 28,75** neu (15 Angebote)**EUR 75,06** gebraucht (3 Angebote)**Elasticsearch: The Definitive Guide** von Clinton Gormley und Zachary Tong von O'Reilly & Assoc Inc. (25. November 2014)

Kostenlose Lieferung möglich und 1 weitere Aktion

**EUR 37,20** Taschenbuch 

Vorbestellbar. Erscheinungsdatum: 25. November 2014

# Suche per Parameter

```
curl -XGET "http://localhost:9200/conferences/talk/_search  
?q=elasticsearch&pretty=true"  
{"took" : 73,  
 [...]  
 "hits" : {  
 [...]  
 "hits" : [ {  
 [...]  
 "_score" : 0.076713204,  
 "_source":{  
 "title" : "Anwendungsfälle für Elasticsearch",  
 "tags" : ["Java", "Lucene"],  
 [...]  
 } } ]  
 }  
 }
```

# Query-DSL

```
curl -XGET "http://localhost:9200/conference/talk/_search" -d'  
{  
  "query": {  
    "filtered": {  
      "query": {  
        "match": {  
          "title": "elasticsearch"  
        }  
      },  
      "filter": {  
        "term": {  
          "tags": "java"  
        }  
      }  
    }  
  }'  
}'
```

# Sprache

```
curl -XGET "http://localhost:9200/conferences/talk/_search  
?q=title:anwendungsfall&pretty=true"  
{  
[...]  
,  
  "hits" : {  
    "total" : 0,  
    "max_score" : null,  
    "hits" : [ ]  
  }  
}
```

# Mapping

```
curl -XDELETE "http://localhost:9200/conferences/"  
  
curl -XPUT "http://localhost:9200/conferences/"  
  
curl -XPUT "http://localhost:9200/conferences/talk/_mapping" -d'  
{  
  "properties": {  
    "title": {  
      "type": "string",  
      "analyzer": "german"  
    }  
  }  
}'
```

# Sprache

```
curl -XGET  
"http://localhost:9200/conferences/talk/_search  
?q=title:anwendungsfall&pretty=true"  
  
{  
[...]  
,  
  "hits" : {  
    "total" : 1,  
    [...]  
  }  
}
```

# Was noch?

- Faceting/Aggregations
- Suggestions
- Highlighting
- Sortierung
- Pagination
- ...

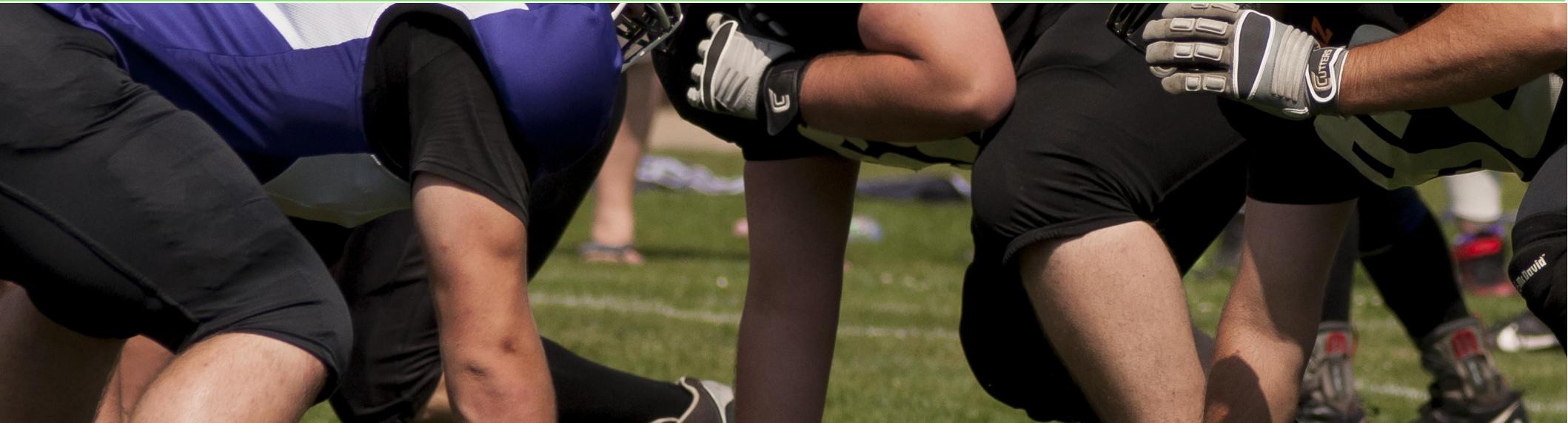
# Recap

- Ausdrucksstarke Suchen über Query DSL
- Analyzing und Mapping
- Lucene bietet viele Search-Features

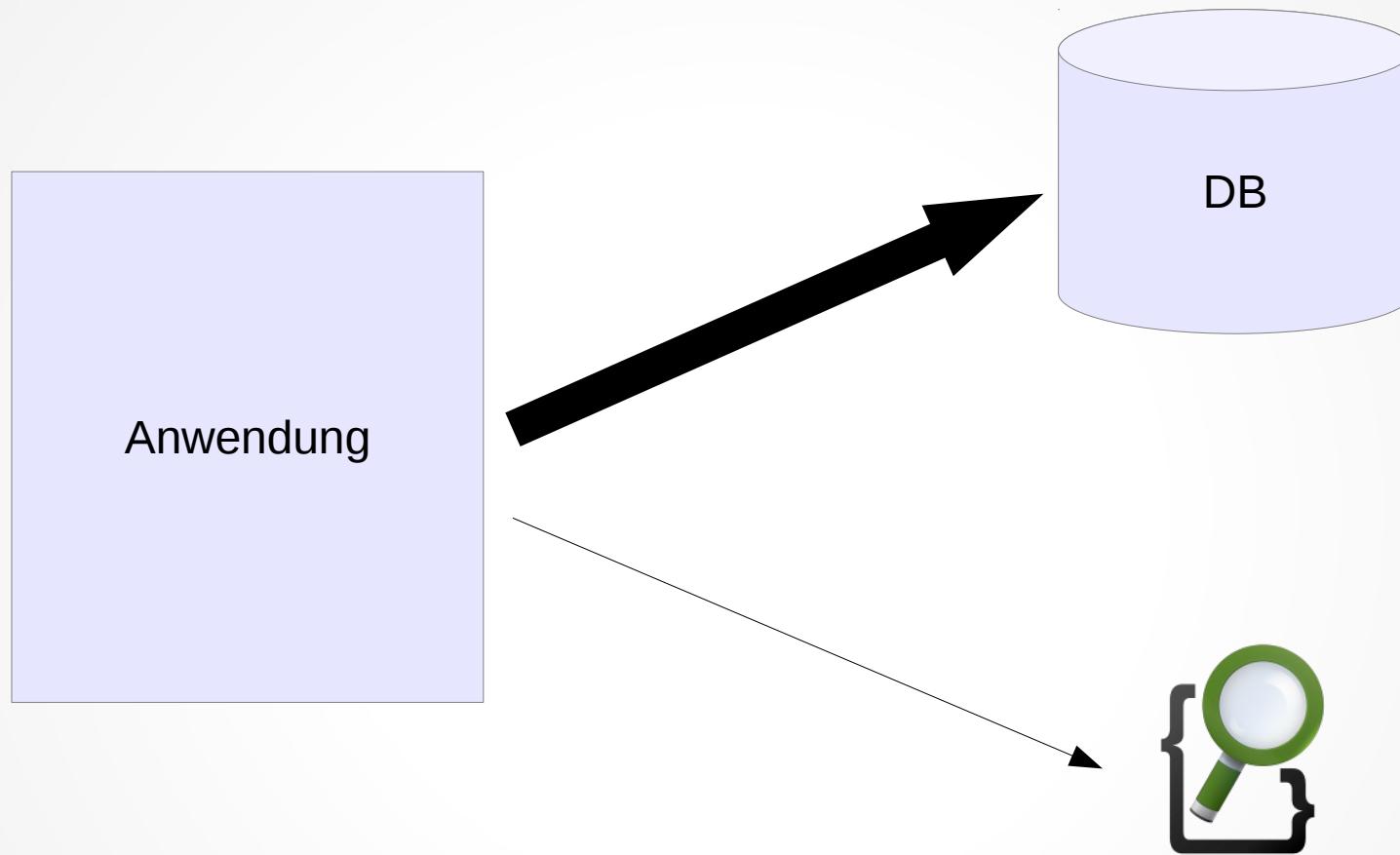
# Users

- GitHub
  - [http://exploringelasticsearch.com/github\\_interview.html](http://exploringelasticsearch.com/github_interview.html)
  - <http://www.elasticsearch.org/case-study/github/>
- StackOverflow
  - <http://meta.stackexchange.com/questions/160100/a-new-search-engine-for-stack-exchange>
  - <http://nickcraver.com/blog/2013/11/22/what-it-takes-to-run-stack-overflow/>
- XING
  - <http://www.elasticsearch.org/case-study/xing/>

# Flexible Cache



# Setup Suche



# Nur Suche?

← → C [www.amazon.de/s/ref=sr\\_nr\\_n\\_5?r=h=n%3A186606%2Cp\\_n\\_publication\\_date%3A1778535031%2Cn%3A%21541686%2Cn%3A124&bbn=541686&ie=UTF8](http://www.amazon.de/s/ref=sr_nr_n_5?r=h=n%3A186606%2Cp_n_publication_date%3A1778535031%2Cn%3A%21541686%2Cn%3A124&bbn=541686&ie=UTF8)

**amazon.de** Mein Amazon Angebote Gutscheine Verkaufen Hilfe

Alle Kategorien ▾ Suche Computer & Internet ▾ Los

Bücher Erweiterte Suche Stöbern Bestseller Neuheiten Hörbücher Fremdsprachige Bücher Taschenbücher Fachbücher Schulbücher Angebote Trade-In

1-16 von 211 Ergebnissen in Bücher : Letzte 30 Tage : Computer & Internet

Ergebnisse anzeigen für

**Neuheiten**

← Auswahl aufheben

**Letzte 30 Tage**

Letzte 90 Tage (686)

Demnächst (527)

← Bücher

**Computer & Internet**

Datenbanken (6)

Grafik & Multimedia (9)

Hardware & Technik (19)

IT-Ausbildung & -Berufe (51)

Mac (5)

Microsoft (18)

Programmierung & Webdesign (24)

Sicherheit (21)

Unix & Linux (3)

Filtern nach

**Format**

Gebundene Ausgabe Taschenbuch Kindle eBook

**Buch-Tipps für Sie**

**Bücher bei Amazon.de**

Entdecken Sie unsere große Auswahl an Büchern und erhalten Sie persönliche Empfehlungen im [Buch-Shop](#).

**Format**

**Ubuntu 14.04 LTS: Aktuell zu »Trusty Tahr [Ziegenartiges]« (Galileo Computing)** von Marcus Fischer von Galileo Computing

**EUR 49,90** Gebundene Ausgabe Prime

Lieferung bis zum **Dienstag, 15. Juli**

Andere Angebote - Gebundene Ausgabe

**EUR 49,90** neu (62 Angebote)

**EUR 49,90** gebraucht (3 Angebote)

Kostenlose Lieferung möglich.

Umtausch ist für Amazon Geschenkgutschein qualifiziert

**Java SE 8 Standard-Bibliothek: Insel 2: Das Handbuch für Java-Entwickler (Galileo Computing)** von Galileo Computing

**EUR 49,90** Gebundene Ausgabe Prime

Lieferung bis zum **Dienstag, 15. Juli**

Andere Angebote - Gebundene Ausgabe

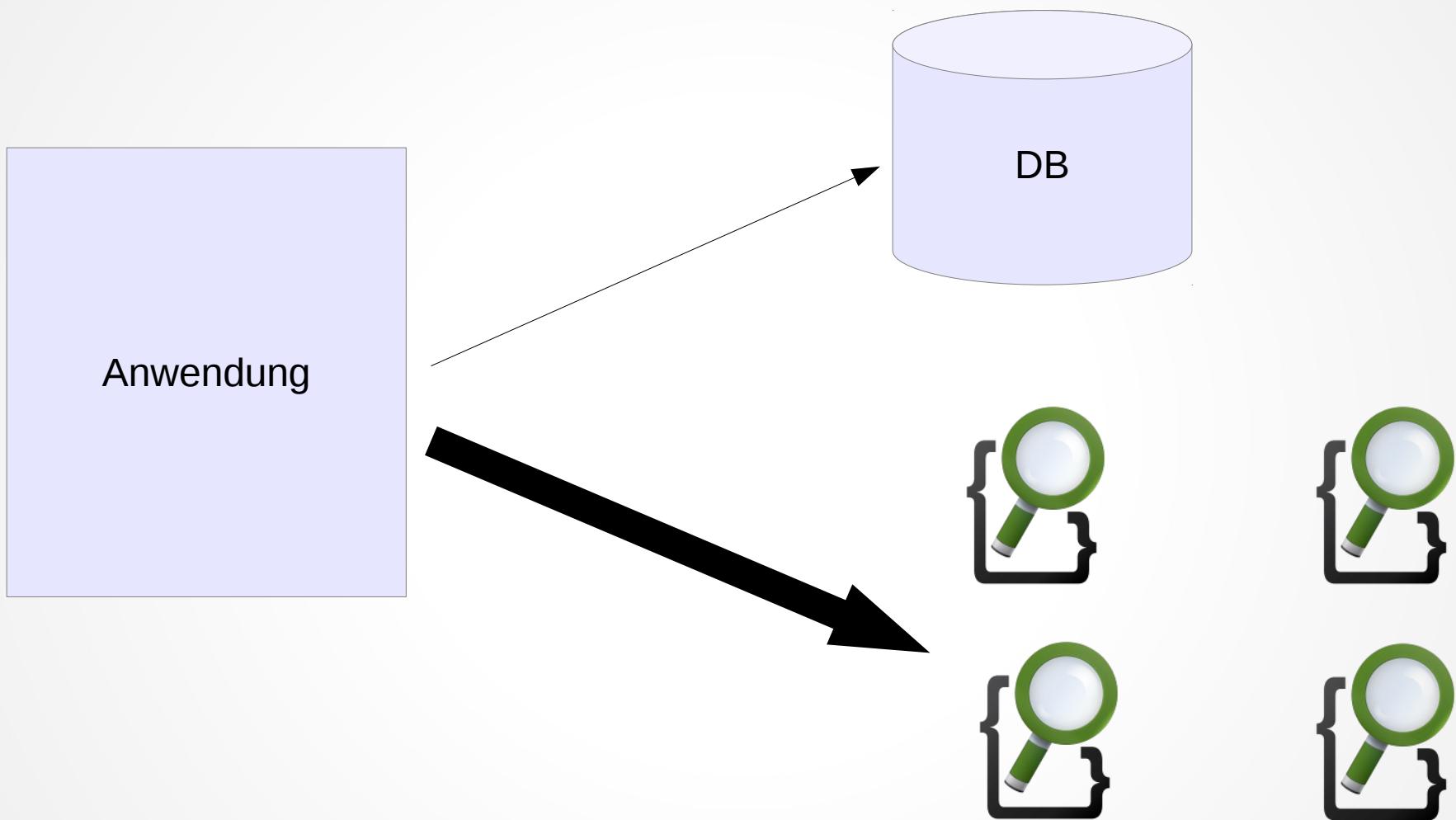
**EUR 49,90** neu (72 Angebote)

**EUR 49,90** gebraucht (3 Angebote)

Kostenlose Lieferung möglich.

Umtausch ist für Amazon Geschenkgutschein qualifiziert

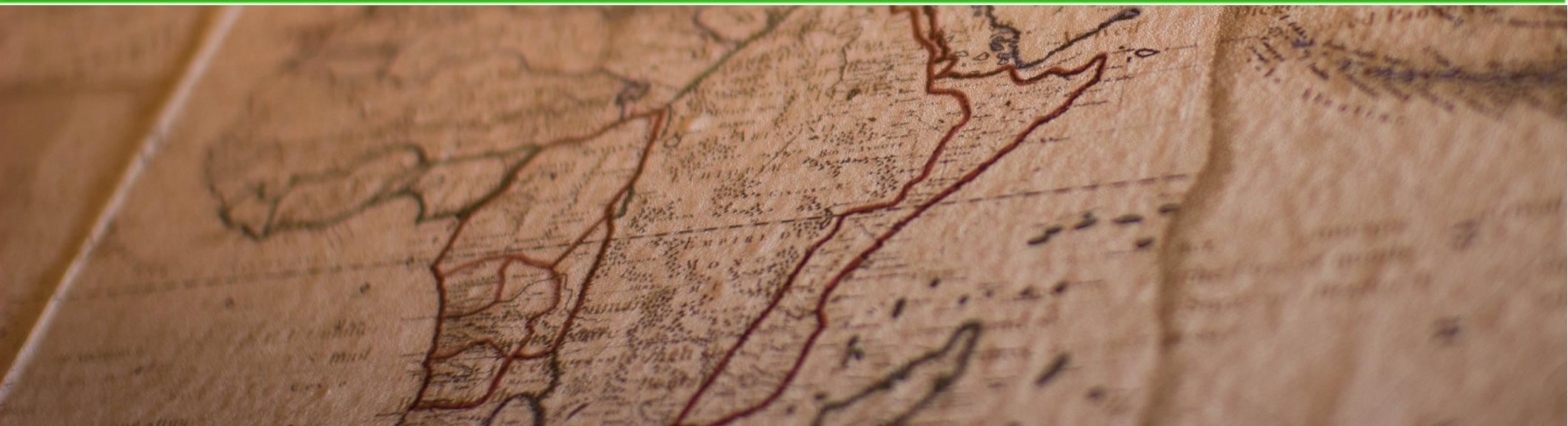
# Queries



# Ergebnislisting

```
curl -XPOST "http://localhost:9200/conferences/_search  
" -d'  
{  
  "filter": {  
    "term": {  
      "tags": "java"  
    }  
  }  
}'
```

# Geo-Suche



# Anwendungen

- Zeige nächste Filiale
- Filialsuche
- Sortierung Kleinanzeigen
- Filterung auf Nähe
- Social Media-Analysen

# Document

```
{  
  "title" : "Anwendungsfälle für Elasticsearch",  
  "speaker" : "Florian Hopf",  
  "date" : "2015-09-15T19:00:00.000Z",  
  "tags" : ["Java", "Lucene"],  
  "conference" : {  
    "name" : "JUG-KL",  
    "city" : "Kaiserslautern",  
    "coordinates": {  
      "lon": "7.752635",  
      "lat": "49.431499"  
    }  
  }  
}
```

# Mapping

```
curl -XPUT "http://localhost:9200/conferences/talk/_mapping" -d '  
{  
  "properties": {  
    [...],  
    "conference": {  
      "type": "object",  
      "properties": {  
        "coordinates": {  
          "type": "geo_point"  
        }  
      }  
    }  
  }'  
}'
```

# Filterung

```
curl -XPOST "http://localhost:9200/conferences/_search" -d'  
{  
  "filter": {  
    "geo_distance": {  
      "conference.coordinates": {  
        "lon": 8.403697,  
        "lat": 49.006616  
      },  
      "distance": "100km",  
      "distance_type": "arc"  
    }  
  }  
}'
```

# Recap

- Elasticsearch kann mehr als Volltext
- Ausgefeilte Geo-Algorithmen
- Sortierung nach Distanz
- Filterung nach Distanz oder Bereich
- Berechnung von Distanz

# User

- Foursquare
  - <http://engineering.foursquare.com/2012/08/09/foursquare-now-uses-elastic-search-and-on-a-related-note-slashem-also-works-with-elasticsearch/>
- Gild
  - <http://www.elasticsearch.org/case-study/gild/>

Datei Bearbeiten Ansicht Lesezeichen Einstellungen Hilfe

```
flo@flo-ThinkPad-T420:~/temp/slides/elasticsearch-1.2.1$ bin/elasticsearch
[2014-06-20 15:05:21,853] [INFO ] [node] [Clive] version[1.2.1], pid[7078], build[6c95b75/2014-06-03T15:02:52Z]
[2014-06-20 15:05:21,853] [INFO ] [node] [Clive] initializing ...
[2014-06-20 15:05:21,858] [INFO ] [plugins] [Clive] loaded [], sites []
[2014-06-20 15:05:23,829] [INFO ] [node] [Clive] initialized
[2014-06-20 15:05:23,829] [INFO ] [node] [Clive] starting ...
[2014-06-20 15:05:23,939] [INFO ] [transport] [Clive] bound_address {inet[/0:0:0:0:0:0:9301]}, publish_address {inet[/172.28.100.209:9301]}
[2014-06-20 15:05:27,069] [INFO ] [cluster.service] [Clive] detected_master [Seth][Uj1ESjAbT4iLLEikTpCR7w][flo-ThinkPad-T420][inet[/172.28.100.209:9300]], reason: zen-disco-receive(from master [[Seth][Uj1ESjAbT4iLLEikTpCR7w][flo-ThinkPad-T420][inet[/172.28.100.209:9300]]])
[2014-06-20 15:05:27,298] [INFO ] [discovery] [Clive] elasticsearch/V7554GKfRSqCB34FVKfvAA
[2014-06-20 15:05:27,309] [INFO ] [http] [Clive] bound_address {inet[/0:0:0:0:0:0:9201]}, publish_address {inet[/172.28.100.209:9201]}
[2014-06-20 15:05:27,312] [INFO ] [node] [Clive] started
[2014-06-20 15:05:33,267] [INFO ] [cluster.service] [Clive] added {[Solo][j4SxMnI7SaaRrzpMYYMoag][flo-ThinkPad-T420][inet[/172.28.100.209:9302]]}, reason: zen-disco-join([Solo][j4SxMnI7SaaRrzpMYYMoag][flo-ThinkPad-T420][inet[/172.28.100.209:9300]])
[2014-06-20 15:05:41,383] [INFO ] [cluster.service] [Clive] removed {[Solo][j4SxMnI7SaaRrzpMYYMoag][flo-ThinkPad-T420][inet[/172.28.100.209:9300]]}
~>[2014-06-20 15:05:44,320] [INFO ] [node] [Clive] stopping
~/runtime/kibana-3.0.0 : python <2>
```

Datei Bearbeiten Ansicht Lesezeichen Einstellungen Hilfe

```
localhost - - [20/Jun/2014 15:03:36] "GET /css/animate.min.css HTTP/1.1" 200 -
```

# Log-Analyse

```
localhost - - [20/Jun/2014 15:03:37] "GET /font/fontawesome-webfont.woff?v=3.2.1 HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:37] "GET /app/dashboards/default.json?1403269417516 HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:37] "GET /app/panels/text/module.html HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:37] "GET /app/panels/text/editor.html HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:37] "GET /app/panels/text/module.js HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:37] "GET /img/kibana.png HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:40] "GET /index.html HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/timepicker/module.html HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/dashboards/blank.json?1403269422113 HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/query/module.html HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/filtering/module.html HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/timepicker/module.js HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/query/module.js HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/timepicker/custom.html HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/query/helpModal.html HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/query/meta.html HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/filtering/module.js HTTP/1.1" 200 -
localhost - - [20/Jun/2014 15:03:42] "GET /app/panels/query/help/lucene.html HTTP/1.1" 200 -
```

```
~/temp/slides/elasticsearch-1.2.1 : java
Datei Bearbeiten Ansicht Lesezeichen Einstellungen Hilfe
[2014-06-20 15:05:07,569] [INFO ] [node]
[2014-06-20 15:05:07,569] [INFO ] [node]
[2014-06-20 15:05:07,573] [INFO ] [plugins]
[2014-06-20 15:05:09,520] [INFO ] [node]
[2014-06-20 15:05:09,520] [INFO ] [node]
```

```
[2014-06-20 15:05:44,418] [INFO ] [cluster.service]
nkPad-T420] [inet[/172.28.100.209:9301]])
```

```
~/temp/slides/elasticsearch-1.2.1 : java
```

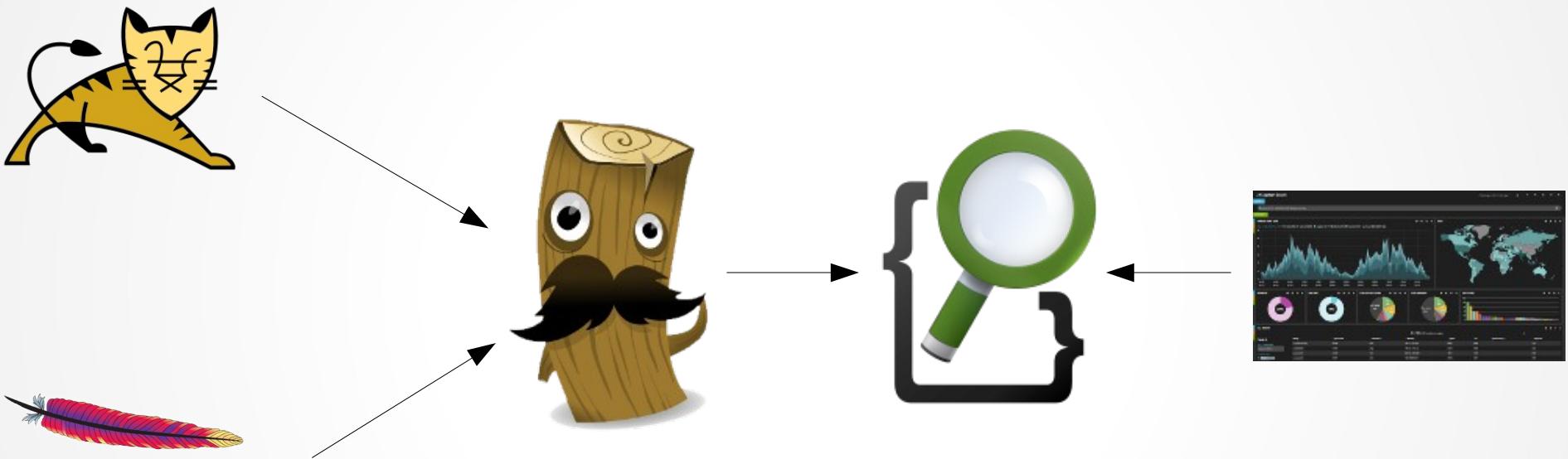
# Logfile-Analyse

- Zentralisierung Logs aus Anwendungen
- Zentralisierung Logs über Maschinen
  - Auch ohne Zugriff
- Leichte Durchsuchbarkeit
- Real-Time-Analysis / Visualisierung
- Daten für alle!

# Logfile-Analyse

- Einlesen
  - Logstash
- Speicherung
  - Elasticsearch
- Auswertung
  - Kibana

# Logfile-Analyse



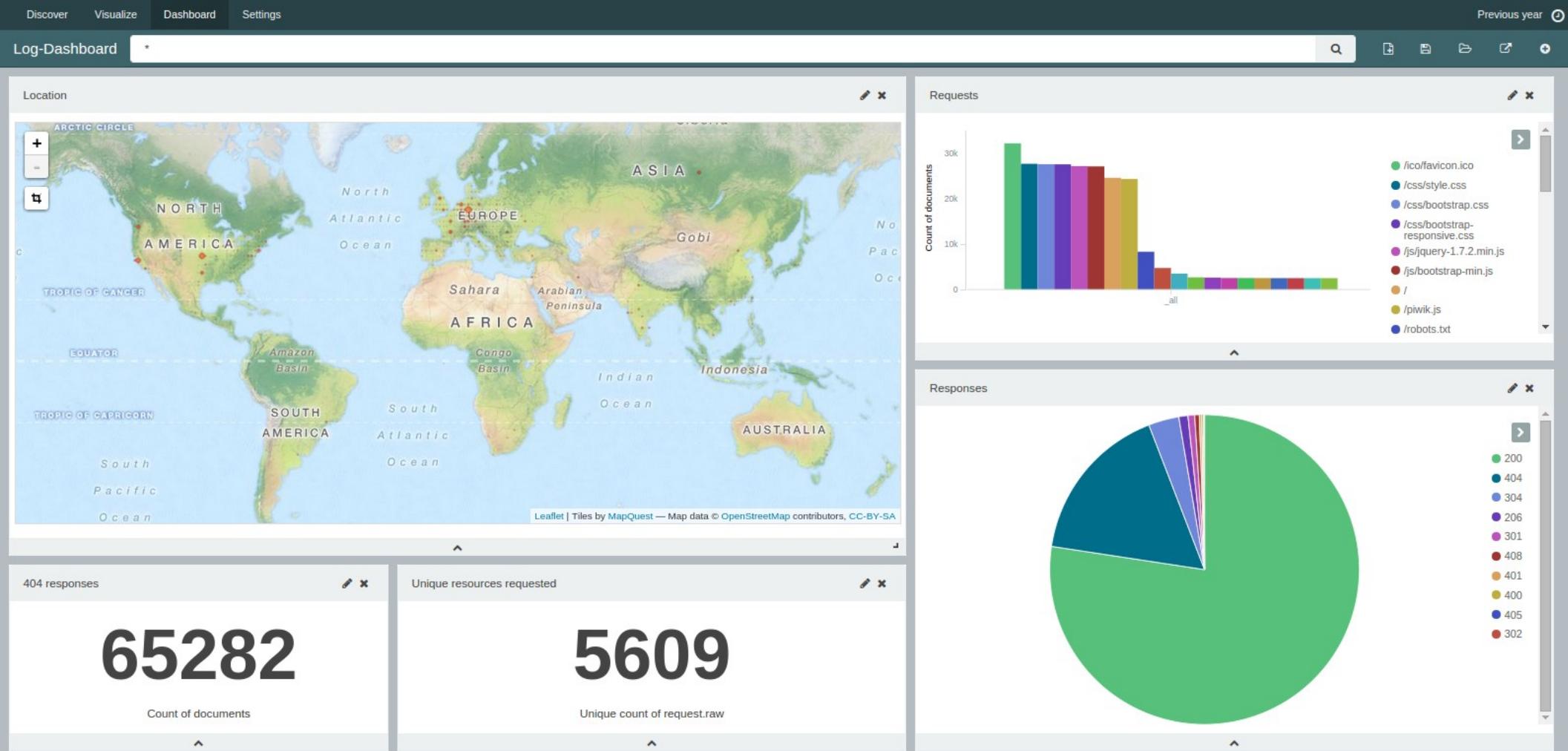
# Logstash-Config

```
input {
  file {
    path => "/var/log/apache2/access.log"
  }
}

filter {
  grok {
    match => { message => "%{COMBINEDAPACHELOG}" }
  }
}

output {
  elasticsearch_http {
    host => "localhost"
  }
}
```

# Kibana



# Recap

- Einlesen, Anreichern, Speichern von Logevents
- Zahlreiche Inputs in Logstash
- Konsolidierung
- Zentralisierung
- Auswertung

# User

- CERN
  - <https://medium.com/@ghoranyi/needle-in-a-haystack-873c97a99983>
- Bloomberg
  - <http://www.elasticsearch.org/videos/using-elasticsearch-logstash-kibana-technologies-centralized-viewing-logs-bloomberg/>



Analytics

# Analytics

- Auswertung auch großer Datenmengen
  - Social Media
  - Data Warehouse
- Aggregationen auf Feldern
- Datenkonsolidierung aus unterschiedlichen Quellen
- Visualisierung

# Aggregations

```
curl -XGET "http://localhost:9200/devoxx/tweet/_search" -d'  
{  
  "aggs" : {  
    "hashtags" : {  
      "terms" : {  
        "field" : "hashtag.text"  
      }  
    }  
  }  
}'
```

# Aggregations

```
"aggregations": {  
    "hashtags": {  
        "buckets": [  
            {  
                "key": "dartlang",  
                "doc_count": 229  
            },  
            {  
                "key": "java",  
                "doc_count": 216  
            },  
            [...]  
        ]  
    }  
}
```

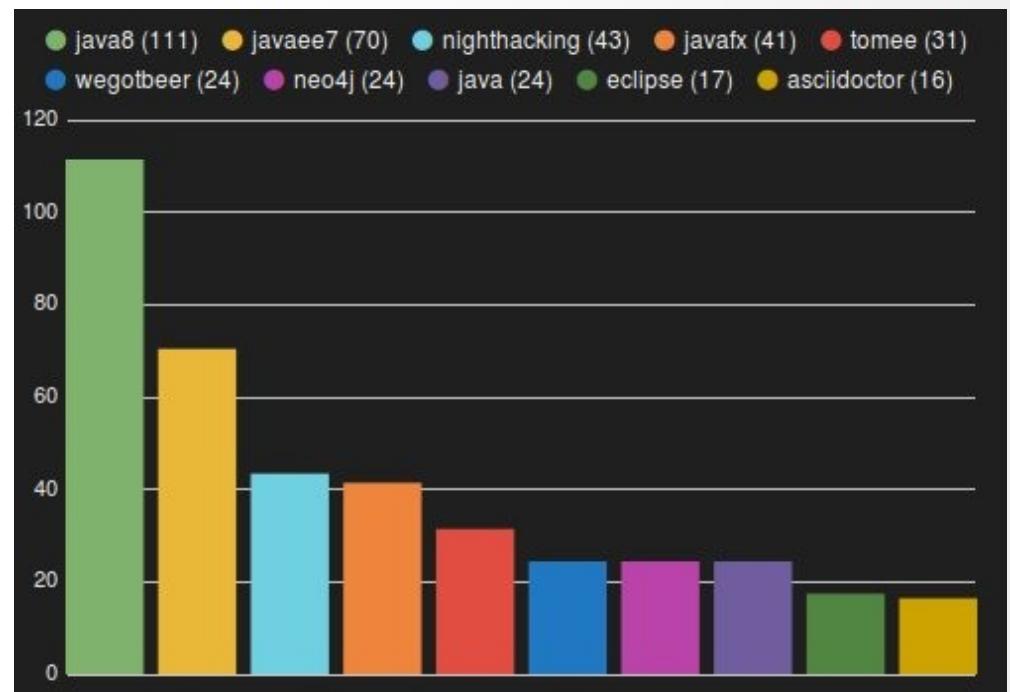
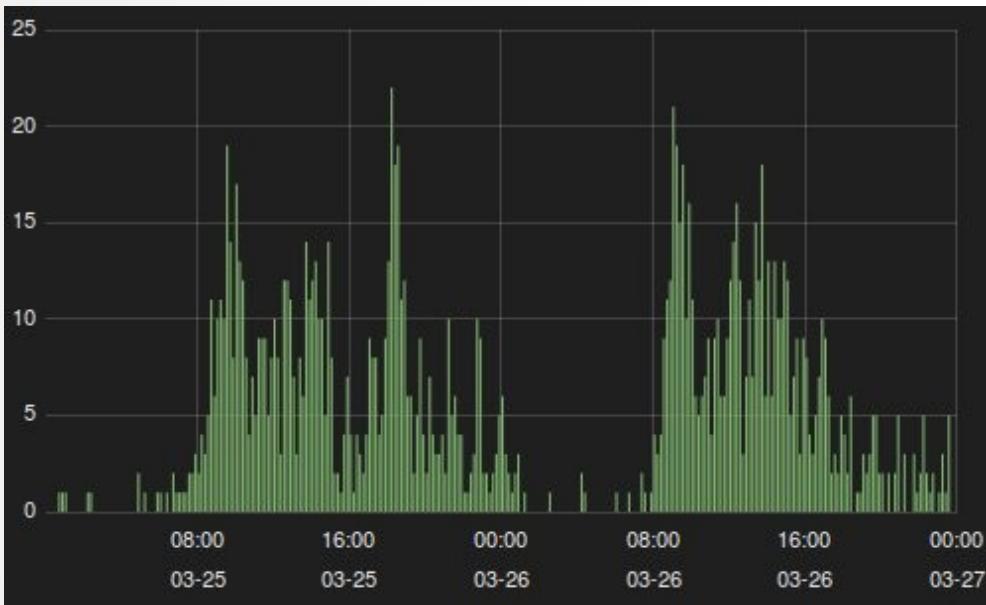
# Aggregations

```
"aggs" : {  
    "hashtags" : {  
        "terms" : {  
            "field" : "hashtag.text"  
        },  
        "aggs" : {  
            "hashtagusers" : {  
                "terms" : {  
                    "field" : "user.screen_name"  
                }  
            }  
        }  
    }  
}
```

# Aggregations

```
"key": "scala",
  "doc_count": 130,
  "hashtagusers": {
    "buckets": [
      {
        "key": "jaceklaskowski",
        "doc_count": 74
      },
      {
        "key": "ManningBooks",
        "doc_count": 3
      },
      [...]
```

# Visualisierung



# Recap

- Auswertung großer Datenmengen
- Visualisierung
- Zahlreiche Aggregationen
  - Bucket-Aggregations
  - Metric-Aggregations

# User

- Cogenta
  - <https://www.elastic.co/use-cases/cogenta>
- The Guardian
  - <http://www.elasticsearch.org/blog/using-elasticsearch-and-logstash-to-serve-billions-of-searchable-events-for-customers/>
  - <http://www.infoq.com/presentations/elasticsearch-guardian>

# Agenda



Document Store



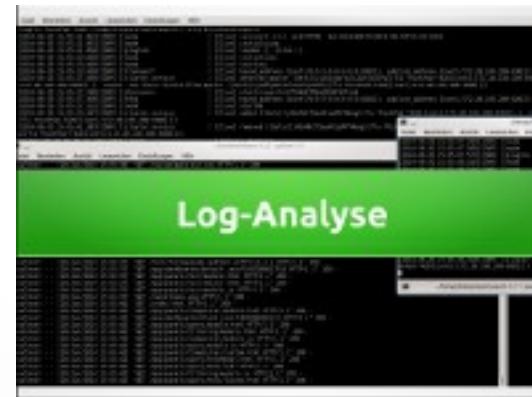
Volltextsuche



Flexibler Cache



Geo-Suche



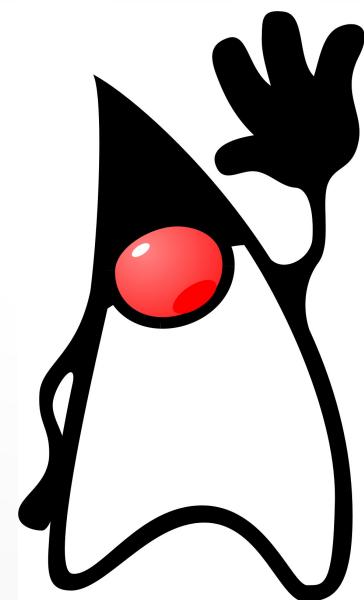
Log-Analyse



Analytics



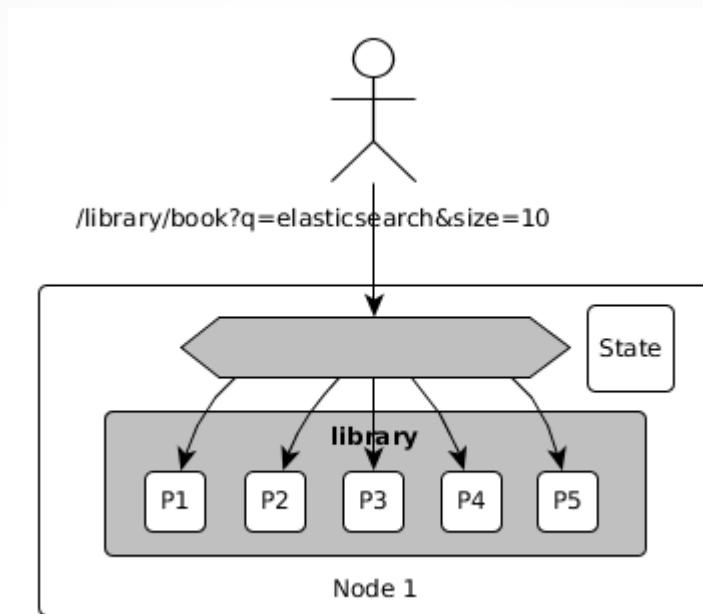
# elasticsearch.



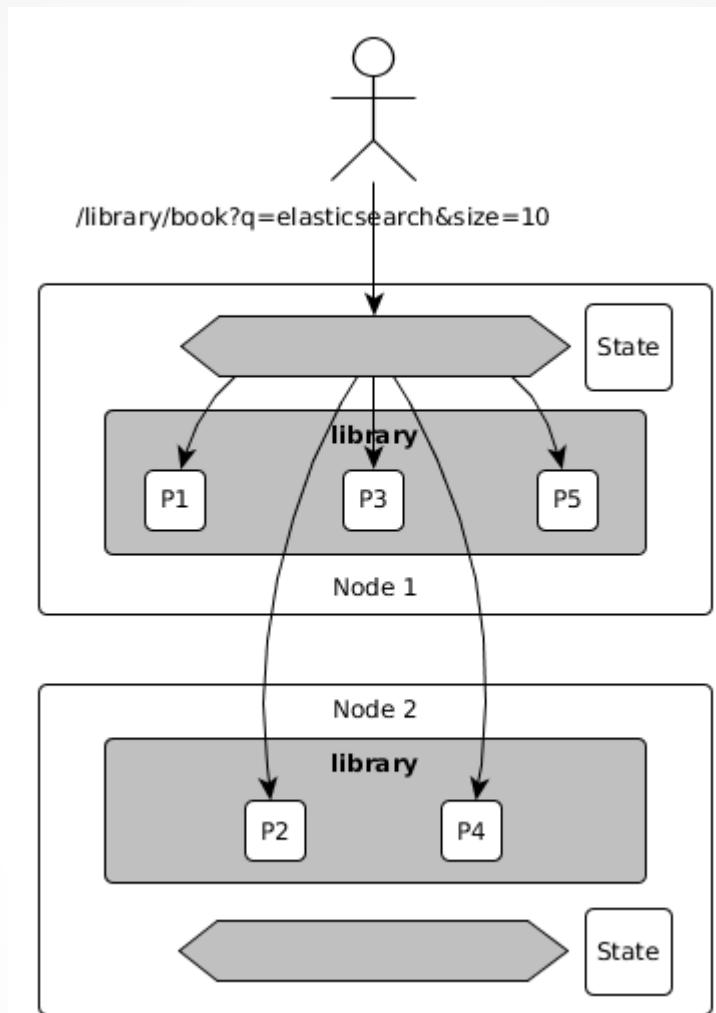
# Verteilung

	library shards: 5 * 2   docs: 1   size: 4.26KB
★ Mangog 172.28.4.45:9301 	
★ Ultron 172.28.4.45:9300 	
★ Unseen 172.28.4.45:9302 	

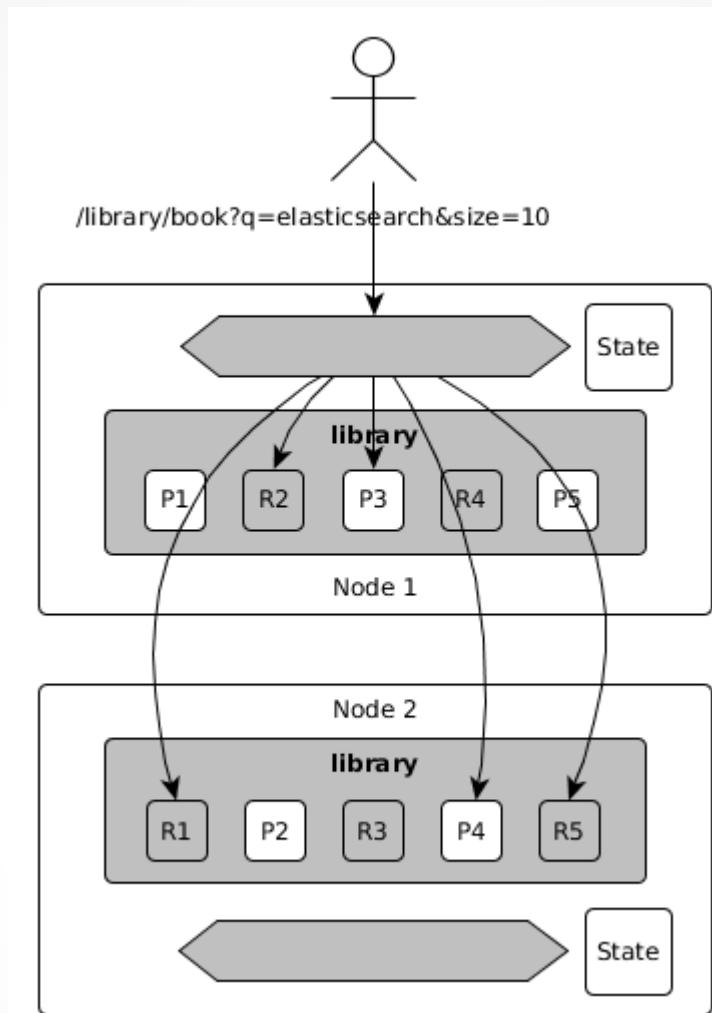
# Suche im Cluster



# Suche im Cluster



# Suche im Cluster



# Standard-Client

# Java!

```
dependencies {  
    compile group: 'org.elasticsearch',  
            name: 'elasticsearch',  
            version: '1.6.0'  
}
```

# Java!

```
TransportAddress address =  
    new InetSocketAddress("localhost", 9300);  
  
Client client = new TransportClient().  
    addTransportAddress(address);
```

# Suche per Query DSL

```
curl -XPOST "http://localhost:9200/library/book/_search"
-d'
{
  "query": {
    "filtered": {
      "query": {
        "match": {
          "title": "elasticsearch"
        }
      },
      "filter": {
        "term": {
          "publisher.name": "manning"
        }
      }
    }
  }
}'
```

# Suche über Java-Client

```
SearchResponse searchResponse = client
    .prepareSearch("library")
    .setQuery(filteredQuery(
        matchQuery("title", "elasticsearch"),
        termFilter("publisher.name", "manning")))
    .execute().actionGet();

assertEquals(1, searchResponse.getHits().getTotalHits());

SearchHit hit = searchResponse.getHits().getAt(0);
assertEquals("Elasticsearch in Action",
    hit.getSource().get("title"));
```

# Indizierung über Java-Client

```
XContentBuilder jsonBuilder = jsonBuilder()  
    .startObject()  
        .field("title", "Elasticsearch")  
        .field("author", "Florian Hopf")  
        .startObject("publisher")  
            .field("name", "dpunkt")  
        .endObject()  
    .endObject();  
  
client.prepareIndex("library", "book")  
    .setSource(jsonBuilder)  
    .execute().actionGet();
```

# Indizierung über Java-Client

```
client.prepareIndex("testindex", "book").|
```

(m) <b>setSource</b> (byte[] source)	IndexRequestBuilder
(m) <b>setSource</b> (byte[] source, int offset, int ...)	IndexRequestBuilder
(m) <b>setSource</b> (BytesReference source)	IndexRequestBuilder
(m) <b>setSource</b> (Map<String, Object> source, XCo...	IndexRequestBuilder
(m) <b>setSource</b> (Object... source)	IndexRequestBuilder
(m) <b>setSource</b> (String field1, Object value1)	IndexRequestBuilder
(m) <b>setSource</b> (String field1, Object value1, S...	IndexRequestBuilder
(m) <b>setSource</b> (String field1, Object value1, S...	IndexRequestBuilder
(m) <b>setSource</b> (String field1, Object value1, S...	IndexRequestBuilder
(m) <b>setSource</b> (String source)	IndexRequestBuilder
(m) <b>setSource</b> (XContentBuilder sourceBuilder)	IndexRequestBuilder
(m) <b>setTimestamper (String timestamper)</b>	IndexRequestBuilder

Dot, semicolon and some other keys will also close this lookup and be inserted into editor

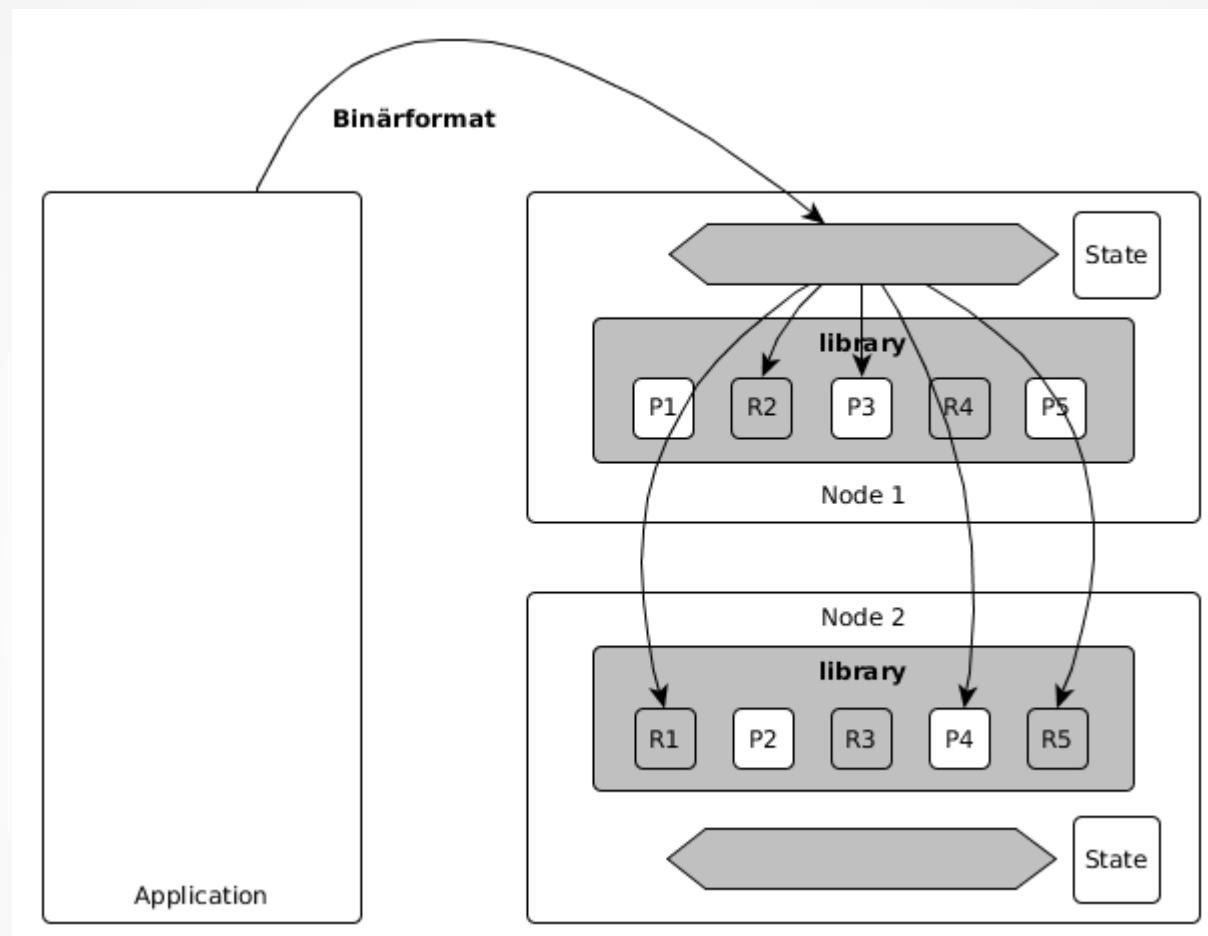
# Transport-Client

- Verbindet sich mit bestehendem Cluster
- Seed-Urls zur Erzeugung

```
TransportAddress address =  
    new InetSocketAddress("localhost", 9300);
```

```
Client client = new TransportClient().  
    addTransportAddress(address);
```

# Transport-Client



# Node-Client

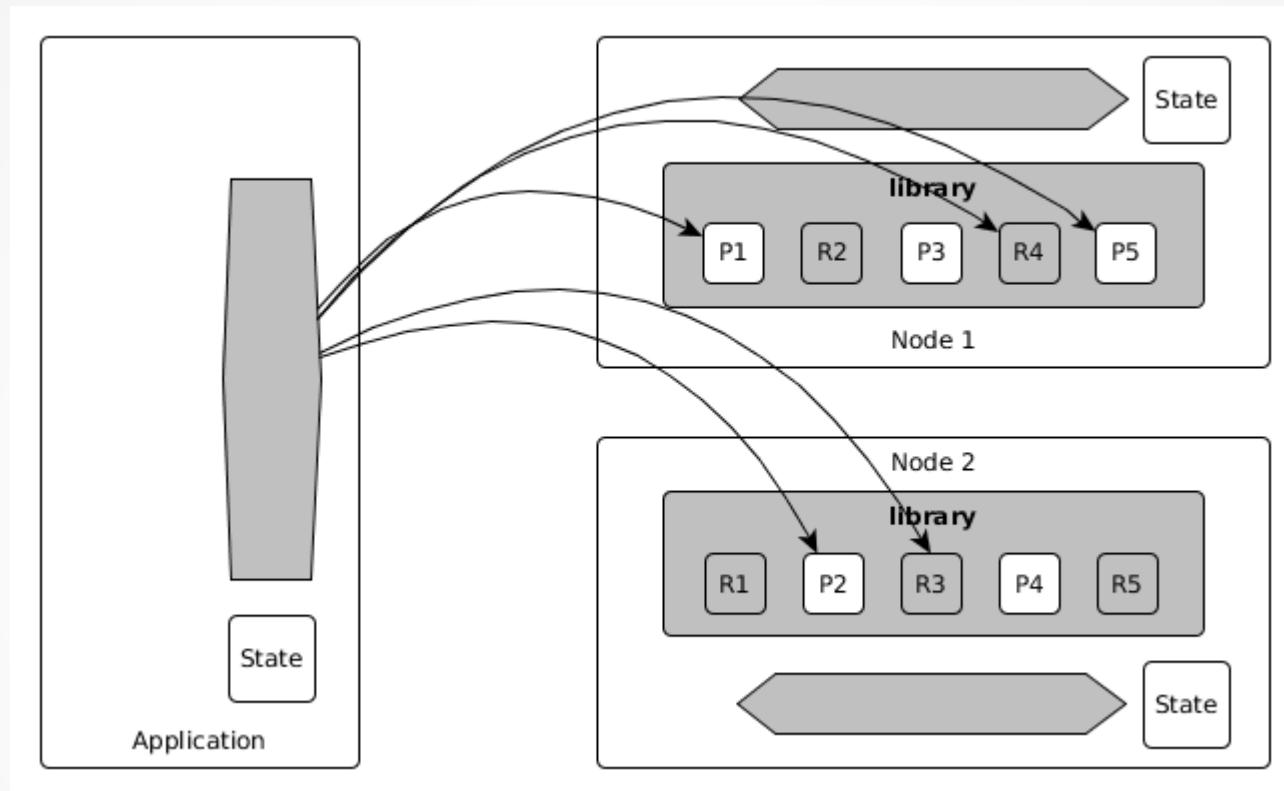
```
Node node = nodeBuilder().client(true).node();  
client client = node.client();
```

# Node-Client

- Startet eigenen Knoten
- Anwendung wird Teil des Clusters

```
Node node = nodeBuilder().client(true).node();  
client client = node.client();
```

# Node-Client



# Standard-Client

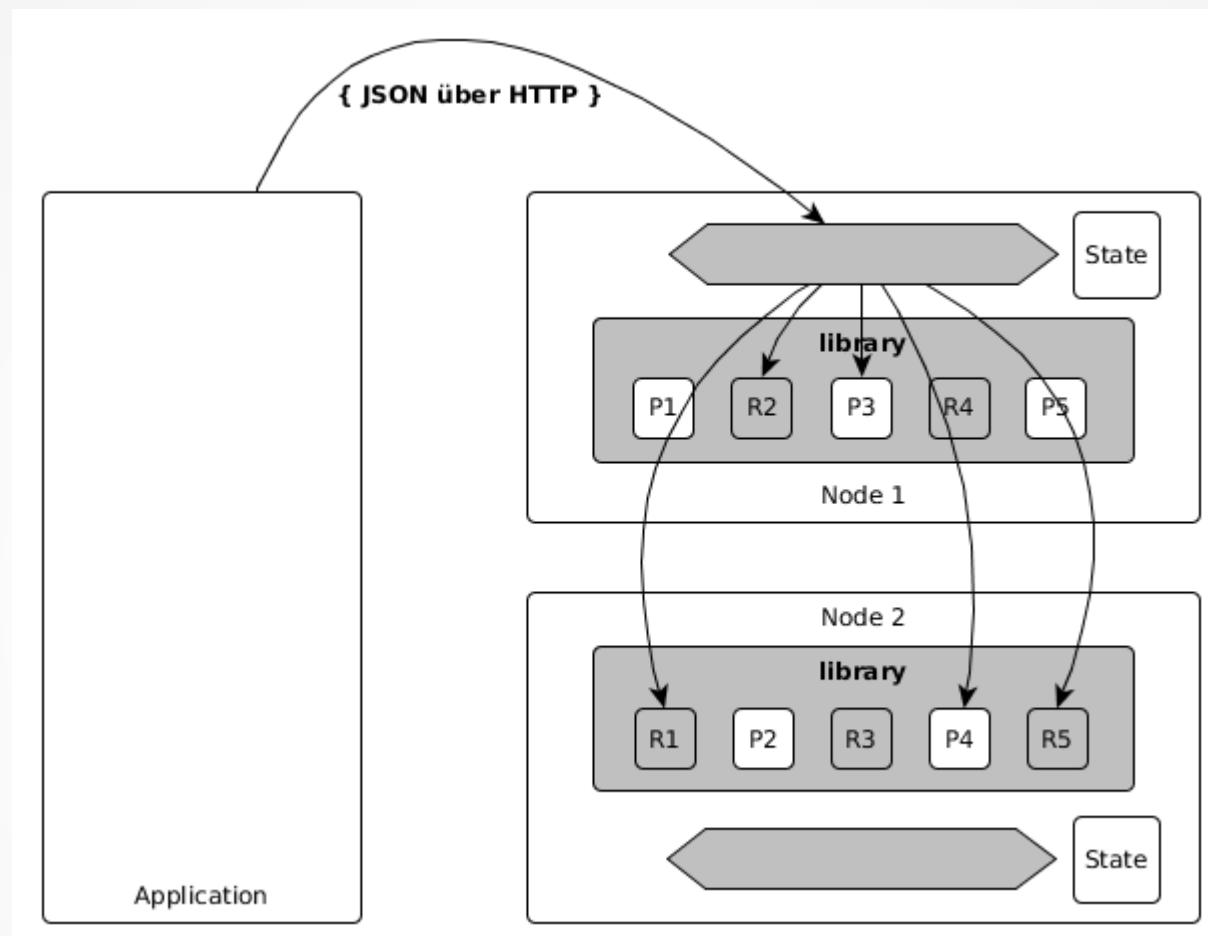
- Volle API-Unterstützung
- Effiziente Kommunikation
- Node-Client
  - Cluster-State spart unnötige Hops

# Standard-Client Gotchas

- Elasticsearch-Version Client == Server
- Node-Client
  - Bei Start und Stop wird Cluster-State übertragen
- Speicherverbrauch
- Elasticsearch-Dependency

# Jest – Http Client

# Jest



# Jest

```
dependencies {  
    compile group: 'io.searchbox',  
            name: 'jest',  
            Version: '0.1.6'  
}
```

# Client

```
JestClientFactory factory = new JestClientFactory();
factory.setHttpClientConfig(new HttpClientConfig
    .Builder("http://localhost:9200")
    .multiThreaded(true)
    .build());
JestClient client = factory.getObject();
```

# Suche mit Jest

```
String query = jsonStringThatMagicallyAppears;  
  
Search search = new Search.Builder(query)  
    .addIndex("library")  
    .build();  
  
SearchResult result = client.execute(search);  
assertEquals(Integer.valueOf(1), result.getTotal());
```

# Suche mit Jest

```
JsonObject jsonObject = result.getJsonObject();
JsonObject hitsObj = jsonObject.getAsJsonObject("hits");
JsonArray hits = hitsObj.getAsJsonArray("hits");
JsonObject hit = hits.get(0).getAsJsonObject();

// ... more boring code
```

# Suche mit Jest

```
public class Book {  
  
    private String title;  
    private List<String> author;  
    private Publisher publisher;  
  
    public Publisher getPublisher() {  
        return publisher;  
    }  
  
    public void setPublisher(Publisher publisher) {  
        this.publisher = publisher;  
    }  
  
    // more code  
}
```

# Suche mit Jest

```
Book book = result.getFirstHit(Book.class).source;  
assertEquals("Elasticsearch in Action", book.getTitle());
```

# Jest

- Alternative HTTP-Implementierung
- Queries als String oder per Elasticsearch-Builder
- Indizieren und Suchen per Java-Beans

# Spring Data Elasticsearch

# Spring Data

- Abstraktionen für Data-Stores
- Spezialitäten bleiben erhalten
- Dynamische Repository-Implementierung
- Populäre Implementierungen
  - Spring Data JPA
  - Spring Data MongoDB

# Dependency

```
dependencies {  
    compile group: 'org.springframework.data',  
            name: 'spring-data-elasticsearch',  
            version: '1.2.0.RELEASE'  
}
```

# Entity

```
@Document(  
    indexName = "library",  
    type = "book")  
public class Book {  
  
    private String id;  
    private String title;  
    private List<String> author;  
    private Publisher publisher;  
  
    // more code  
}
```

# Repository

```
public interface BookRepository  
    extends ElasticsearchCrudRepository<Book, String> {  
  
    public Iterable<Book> findByTitle(String title);  
  
}
```

# Configuration

```
<elasticsearch:node-client id="client" />

<bean name="elasticsearchTemplate"
  class="o.s.d.elasticsearch.core.ElasticsearchTemplate">
  <constructor-arg name="client" ref="client"/>
</bean>

<elasticsearch:repositories
  base-package="de.fhopf.elasticsearch.springdata" />
```

# Indexing

```
Book book = new Book();
book.setTitle("Elasticsearch");
book.setAuthor(Arrays.asList("Florian Hopf"));
book.setPublisher(new Publisher("dpunkt"));

repository.save(book);
```

# Searching

```
Iterable<Book> books = repository.findAll();  
books = repository.findByTitle("Elasticsearch");
```

# Recap

- Abstraktion über Elasticsearch
- Entity-Beans
- Dynamische Repository-Implementierung
- Junges Projekt

# Recap

- Standard-Client
  - Node- oder Transport-Client
- Jest
  - Http-Client mit Java-Bean-Unterstützung
- Spring-Data-Elasticsearch
  - Annotationzentrierte Konfiguration, dynamische Repository-Implementierungen

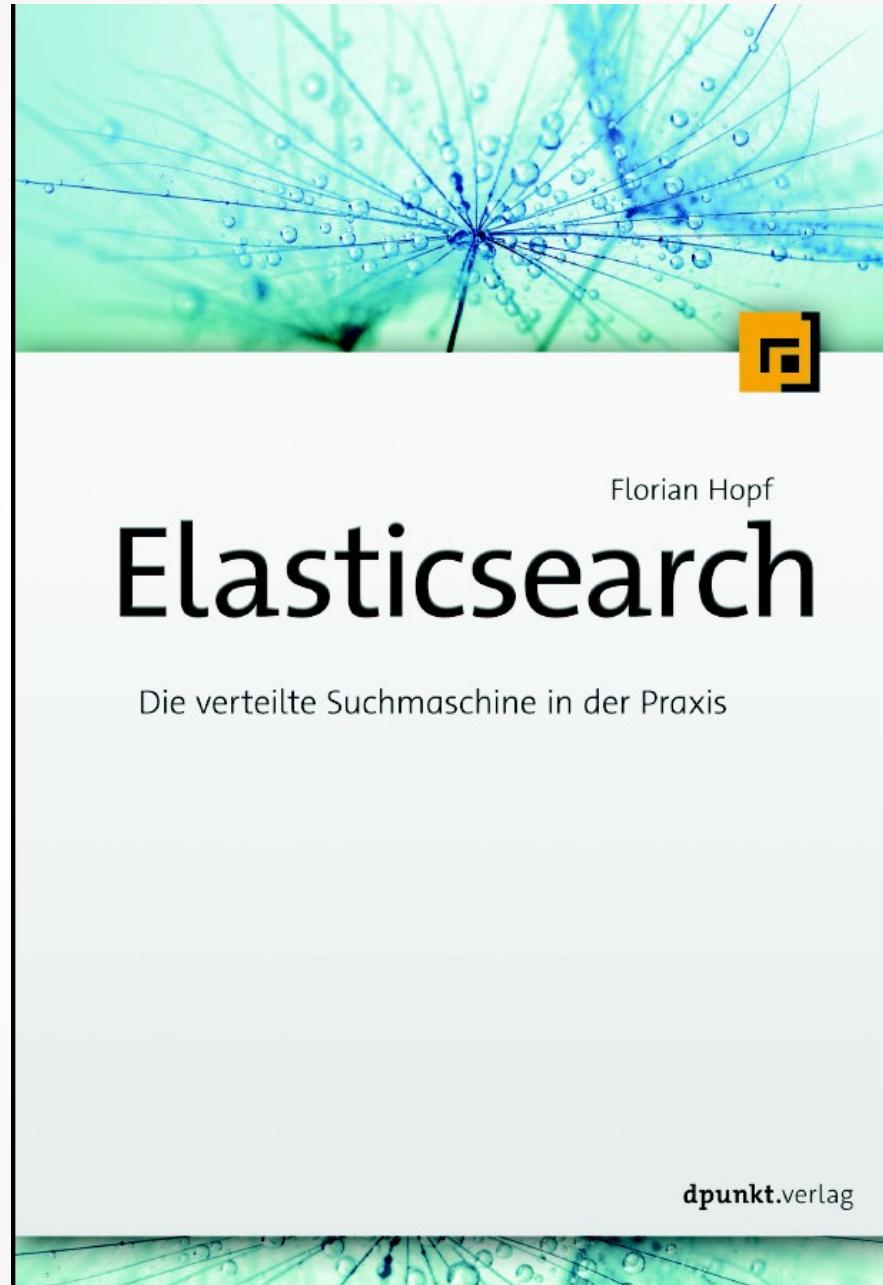
# JVM?

- Clojure: Elastisch
  - Standard-Client oder HTTP
- Groovy: elasticsearch-groovy
  - Standard-Client, Closure-Unterstützung
- Scala: große Auswahl, z.B. elastic4s
  - Standard-Client, Type-Safety

# Links

- <https://www.found.no/foundation/java-clients-for-elasticsearch/>
- <http://elastic.co>
- <https://github.com/searchbox-io/Jest>
- <https://github.com/spring-projects/spring-data-elasticsearch>
- <http://www.florian-hopf.de>

# Weitere Infos



A photograph of a long, straight asphalt road stretching towards a range of snow-capped mountains under a clear blue sky. The road has yellow dashed lines in the center and white solid lines on the sides. The surrounding landscape is flat and green, with some fields and fences on either side.

# Vielen Dank!

@fhopf  
mail@florian-hopf.de  
<http://blog.florian-hopf.de>

# Images

- <http://www.morguefile.com/archive/display/685952>
- <http://www.morguefile.com/archive/display/2359>
- <http://www.morguefile.com/archive/display/615356>
- <http://www.morguefile.com/archive/display/914733>
- <http://www.morguefile.com/archive/display/826258>
- <http://www.morguefile.com/archive/display/170605>
- <http://www.morguefile.com/archive/display/181488>