

SUCCESS



Zentralisiertes Log Management

Ein Erfahrungsbericht



Michael Mimo Moratti
mimo@mimo.ch

0 to 100 of 500 available for paging

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
access_security	MY:VN	html	167.12.22.189	8540	1068	
access_info	IT:MM	png	164.87.170.73	2045	1903	
access_info	AR:ES	html	222.23.102.238	1801	1133	
access_info	IN:IN	html	138.226.66.81	7029	1801	

Wer bin ich...

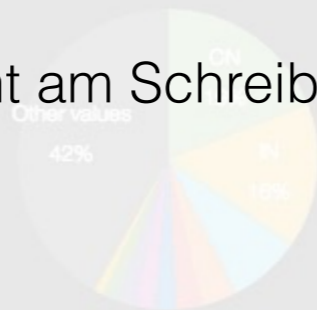
Michael Mimo Moratti, mimo@mimo.ch, jmimo on Github

Java, C, C++, Python, Lua, Bash, Linux, Web zeugs, ...
Architektur, WAF / Entry Server, Code review, Kernel hacking.

3 RLK Installationen aufgebaut

Die Größte Installation: CSS Versicherung (+740 mio. Events)

Und wenn ich gerade mal nicht am Schreibtisch stehe dann:



SUCCESS



Warum sollten wir das tun?



0 to 100 of 500 available for paging

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
access_security	MY:VN	html	167.12.22.189	8540	1066	
access_info	IT:MM	png	164.87.170.73	2045	1903	
access_info	AR:ES	html	222.23.102.238	1801	1133	
access_info	IN:DZ	html	136.226.66.81	7029	1801	



Ich hatte gerade eine Fehlerantwort, was geschah da genau?

Irgendetwas geht nicht, schau doch bitte schnell nach!

Kannst du bitte die Logs von Gestern zwischen 14:36 und 15:12 auf Fehler prüfen!

Kannst du bitte alle Logs der letzten Woche auf Fehler überprüfen!

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
cess_security	MY:VN	html	167.12.22.189	8540	1068	
cess_info	IT:MM	png	164.87.170.73	2045	1903	
cess_info	AR:ES	html	222.23.102.238	1801	1133	
cess_info	IN:IZ	html	138.226.66.81	7029	1801	

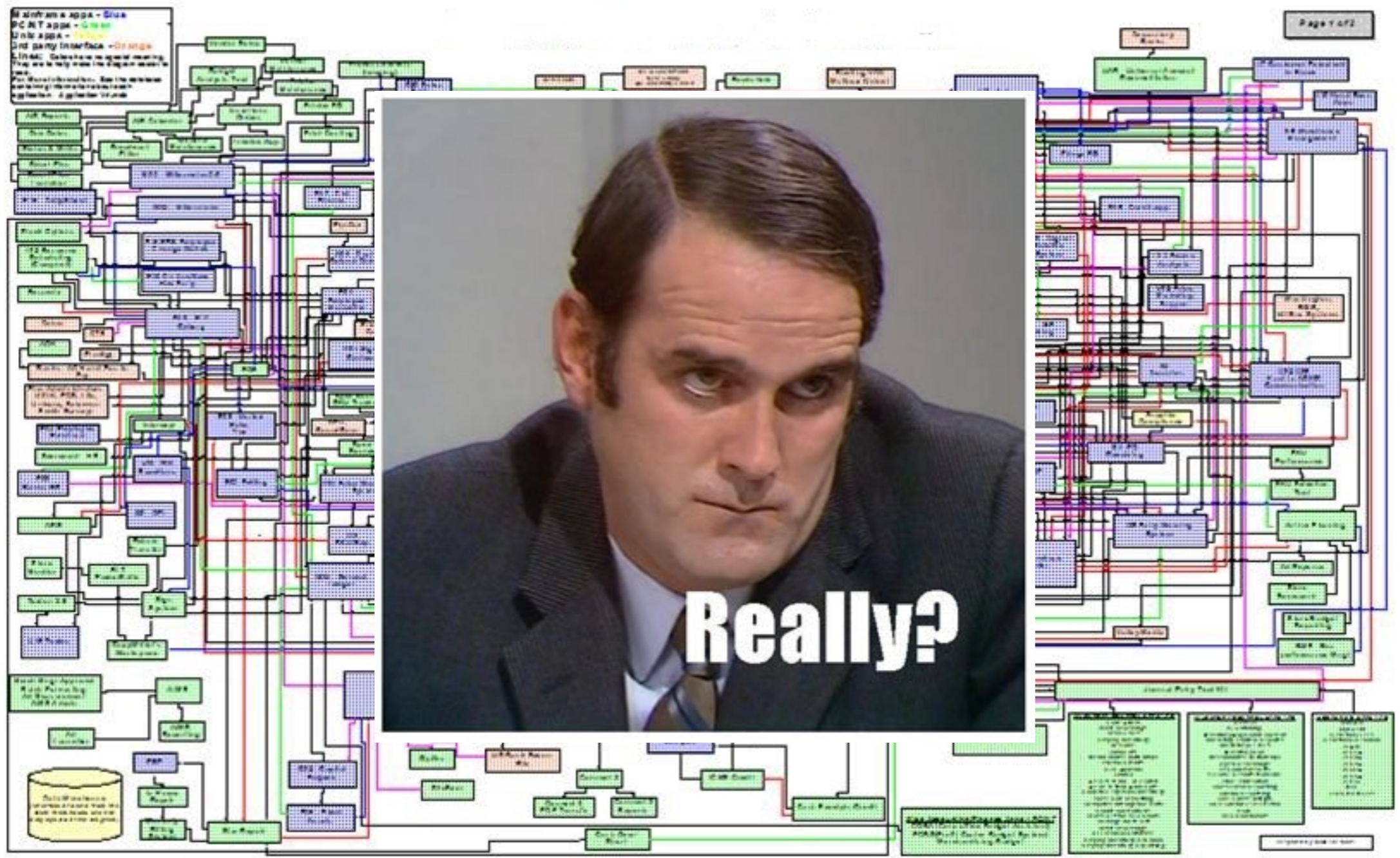
SUCCESS

php (ext) (65

12:00
02-24

+ x V

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
access_security	MY:VN	html	167.12.22.189	8540	1068	
access_info	IT:MM	png	164.87.170.73	2045	1903	
access_info	AR:ES	html	222.23.102.238	1801	1133	
access_info	IN:Z	html	136.226.66.81	7029	1801	



Really?

Und jetzt...

- Werden wir uns an einer Vielzahl von Systemen anmelden.
- Mit “find”, “grep” und anderen Tools die Logfiles nach Anzeichen von Fehler durchsuchen.
- Und versuchen die gefundenen Daten zu Korrelieren und Interpretieren.

```

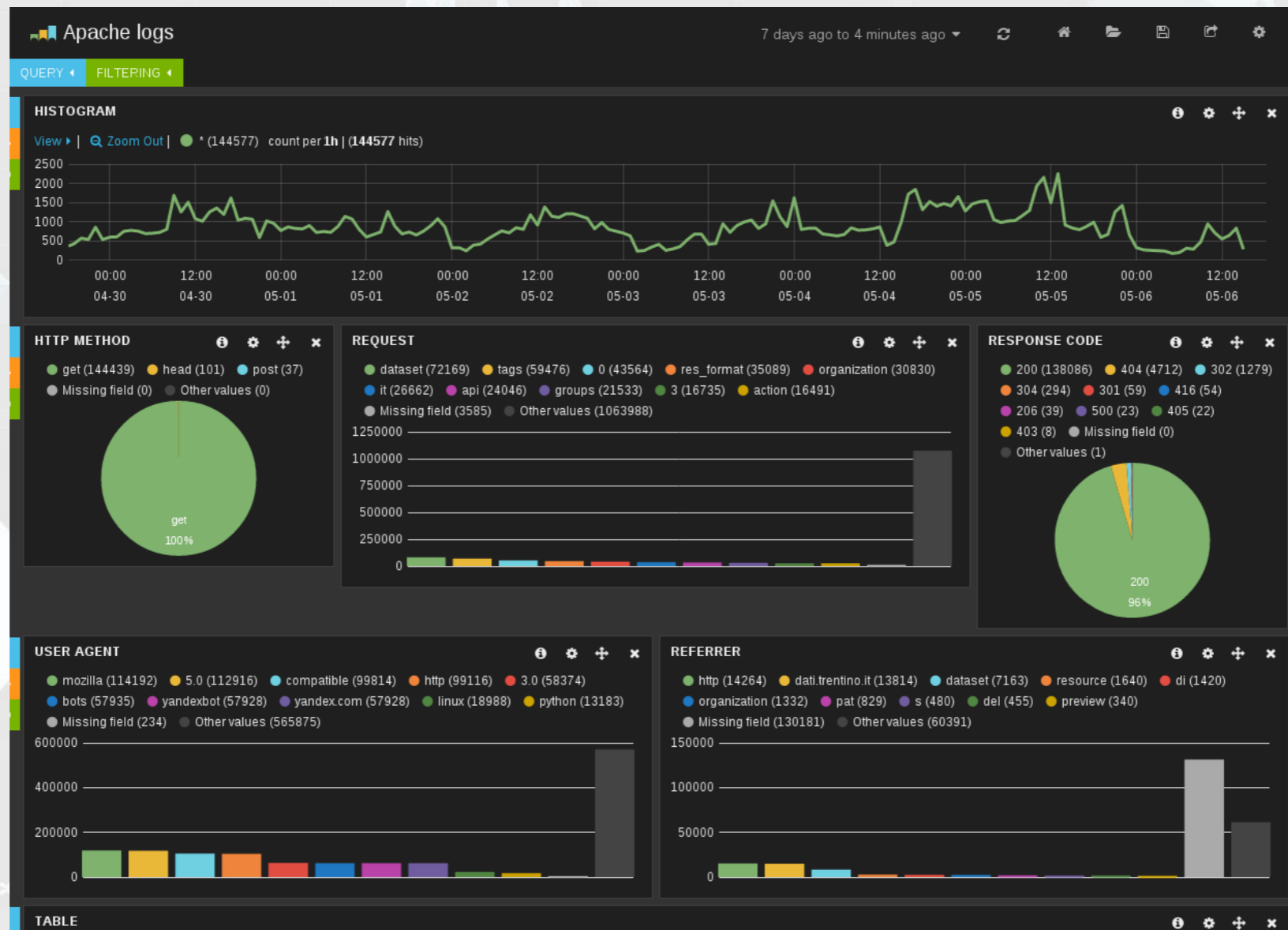
Jan 10 00:51:37 BigThought.local configd[25]: LINKLOCAL en0: parent has no IP
Jan 10 00:51:37 BigThought.local networkd[198]: +[NETLdBelly stopFastFail] Clearing ledbelly failure cache
Jan 10 00:51:37 BigThought.local configd[25]: network changed: v4(en0-:192.168.0.20) DNS- Proxy-
Jan 10 00:51:37 BigThought.local discoveryd[50]: Basic Warn DD_Warn: Corrupt NSEC RDATA size
Jan 10 00:51:37 BigThought.local discoveryd[50]: Basic WABServer NetResolverEvent no resolvers, resetting domains
Jan 10 00:51:37 BigThought.local discoveryd[50]: Basic Bonjour Failed to delete registration recordId=518 error=4
Jan 10 00:51:38 BigThought.local networkd[198]: +[NETLdBelly stopFastFail] Clearing ledbelly failure cache
Jan 10 00:51:38 BigThought.local configd[25]: network changed: v4(en0+:192.168.0.20) DNS+ Proxy+ SMB
Jan 10 00:51:38 BigThought.local discoveryd[50]: Basic Sockets,Warn UDS FD=90 ERROR: Send failed errno=32
Jan 10 00:51:39 BigThought.local discoveryd[50]: Basic Bonjour Failed to delete registration recordId=522 error=4
Jan 10 00:51:46 BigThought.local sharingd[7192]: 00:51:46.555 : SDStatusMonitor::kStatusWirelessPowerChanged
Jan 10 00:51:46 BigThought.local sharingd[302]: 00:51:46.555 : SDStatusMonitor::kStatusWirelessPowerChanged
Jan 10 00:51:46 BigThought.local sharingd[7192]: 00:51:46.604 : SDStatusMonitor::kStatusWirelessPowerChanged
Jan 10 00:51:46 BigThought.local sharingd[302]: 00:51:46.604 : SDStatusMonitor::kStatusWirelessPowerChanged
Jan 10 00:51:47 BigThought.local ntpd[199]: wake time set +0.335678 s
Jan 10 00:51:55 BigThought.local configd[25]: [0x7fe689e54af0] [m]DNS query timeout (query time = 13.752251), [46TE]
Jan 10 00:52:10 BigThought.local configd[25]: [0x7fe689d43ee0] [m]DNS query timeout (query time = 29.040382), [46TE]
Jan 10 00:52:13 BigThought.local configd[25]: [0x7fe689d42920] [m]DNS query timeout (query time = 32.261717), [46TE]

```

process	geo.src	extension	client	bytes	id
process,security	MY:VN	html	167.12.22.189	8540	1068
process,info	IT:MM	png	164.87.170.73	2045	1903
process,info	AR:ES	html	222.23.102.238	1801	1133
process,info	IN:IZ	html	136.226.66.81	7029	1801

Aber wirklich wollen wir...

Eine Analyse über alle Logdaten aller Systeme mit Hilfe einer Query und Filter basierten Visualisierung im Web-browser machen.



Nur, was braucht es dazu...



Klassisches (manuelles) Vorgehen



Sammeln

(scp / ftp ...)



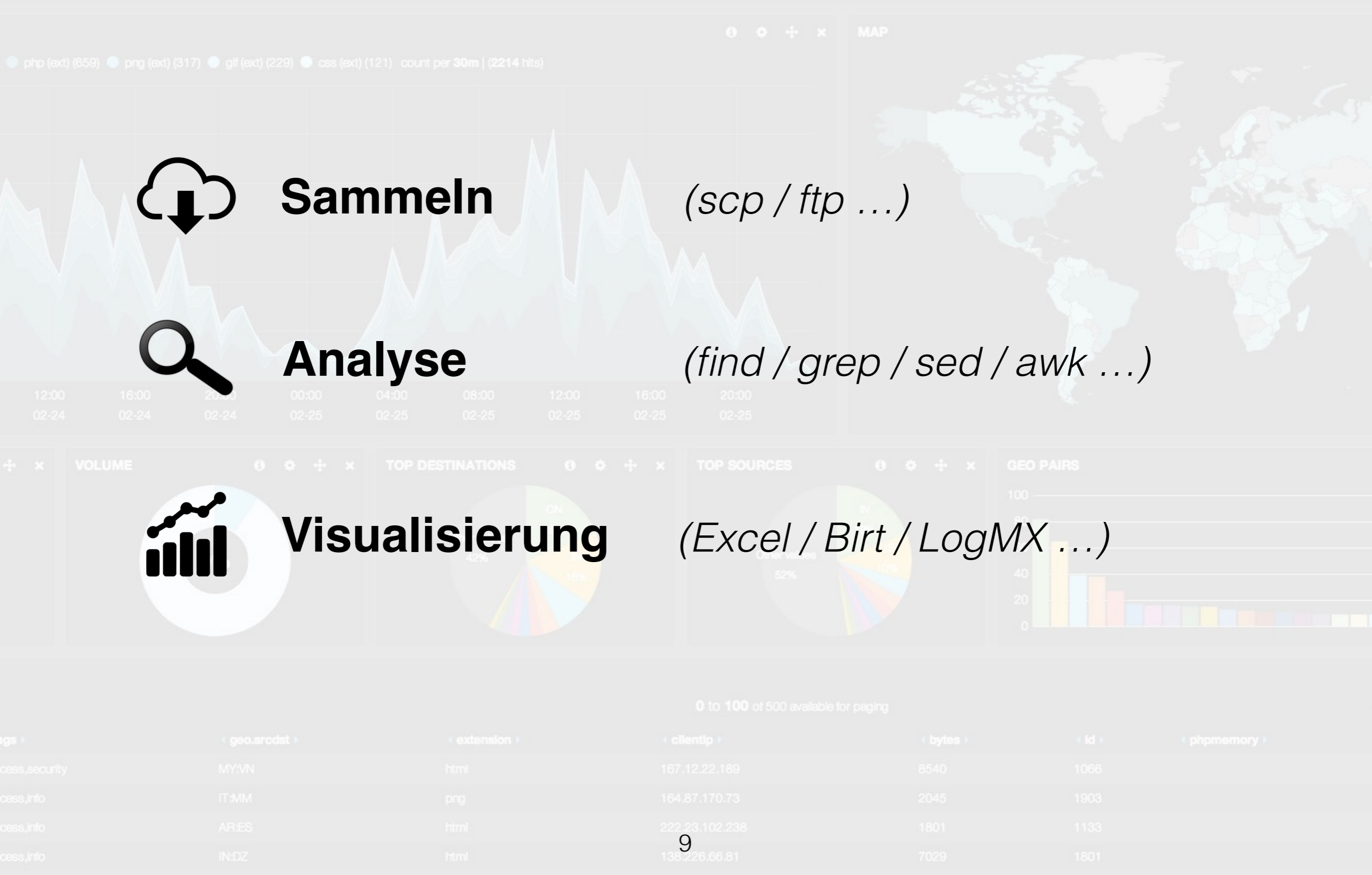
Analyse

(find / grep / sed / awk ...)



Visualisierung

(Excel / Birt / LogMX ...)



ELK Stack



0 to 100 of 500 available for paging

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
access_security	MY:VN	html	167.12.22.189	8540	1068	
access_info	IT:MM	png	164.87.170.73	2045	1903	
access_info	AR:ES	html	222.23.102.238	1801	1133	
access_info	IN:DZ	html	138.226.66.81	7029	1801	



elasticsearch.

Elasticsearch

- NoSQL “Big data” Volltext Suchmaschine
- Java / basiert auf “Apache Lucene”
- Suchen und Indexieren
- Verteilt (shards & copies)
- Clustering
- API — JSON / RESTful



0 to 100 of 500 available for paging

geo.srcdst	extension	clientip	bytes	id	phpmemory
MY:VN	html	167.12.22.189	8540	1068	
IT:MM	png	164.87.170.73	2045	1903	
AR:ES	html	222.23.102.238	1801	1133	
IN:IZ	html	138.226.66.81	7029	1801	



elasticsearch.

Elasticsearch

SUCCESS

php (ext) (659) png (ext) (317) gif (ext) (229) css (ext) (121) count per 30m | (2214 hits)

Elasticsearch Cluster

Node 1

Shard 1

Replica 2

Shard 3

Replica 4

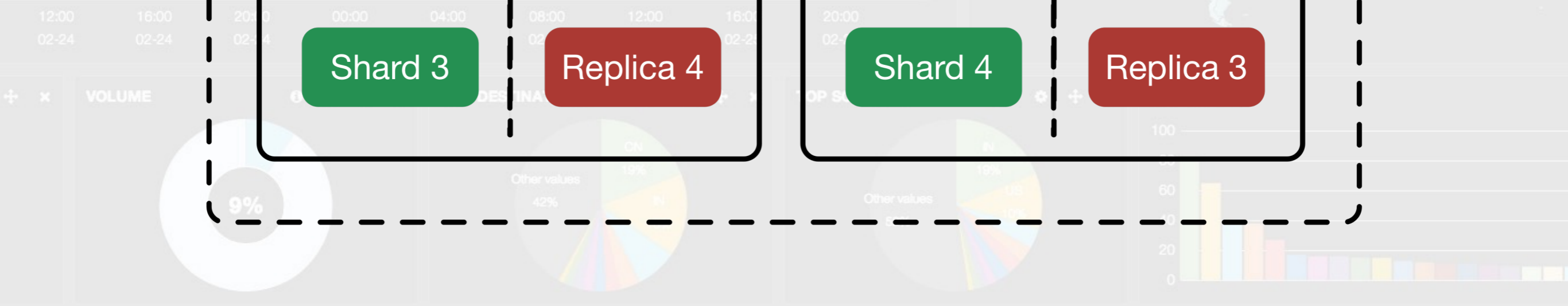
Node 2

Shard 2

Replica 1

Shard 4

Replica 3



0 to 100 of 500 available for paging

geo.srcdst	extension	clientip	bytes	id	phpmemory
MY:VN	html	167.12.22.189	8540	1068	
IT:MM	png	164.87.170.73	2045	1903	
AR:ES	html	222.23.102.238	1801	1133	
IN:Z	html	138.226.66.81	7029	1801	



Logstash

• “Schweizer Sackmesser” zum Sammeln, Transformieren und Weiterleiten von Logdaten.

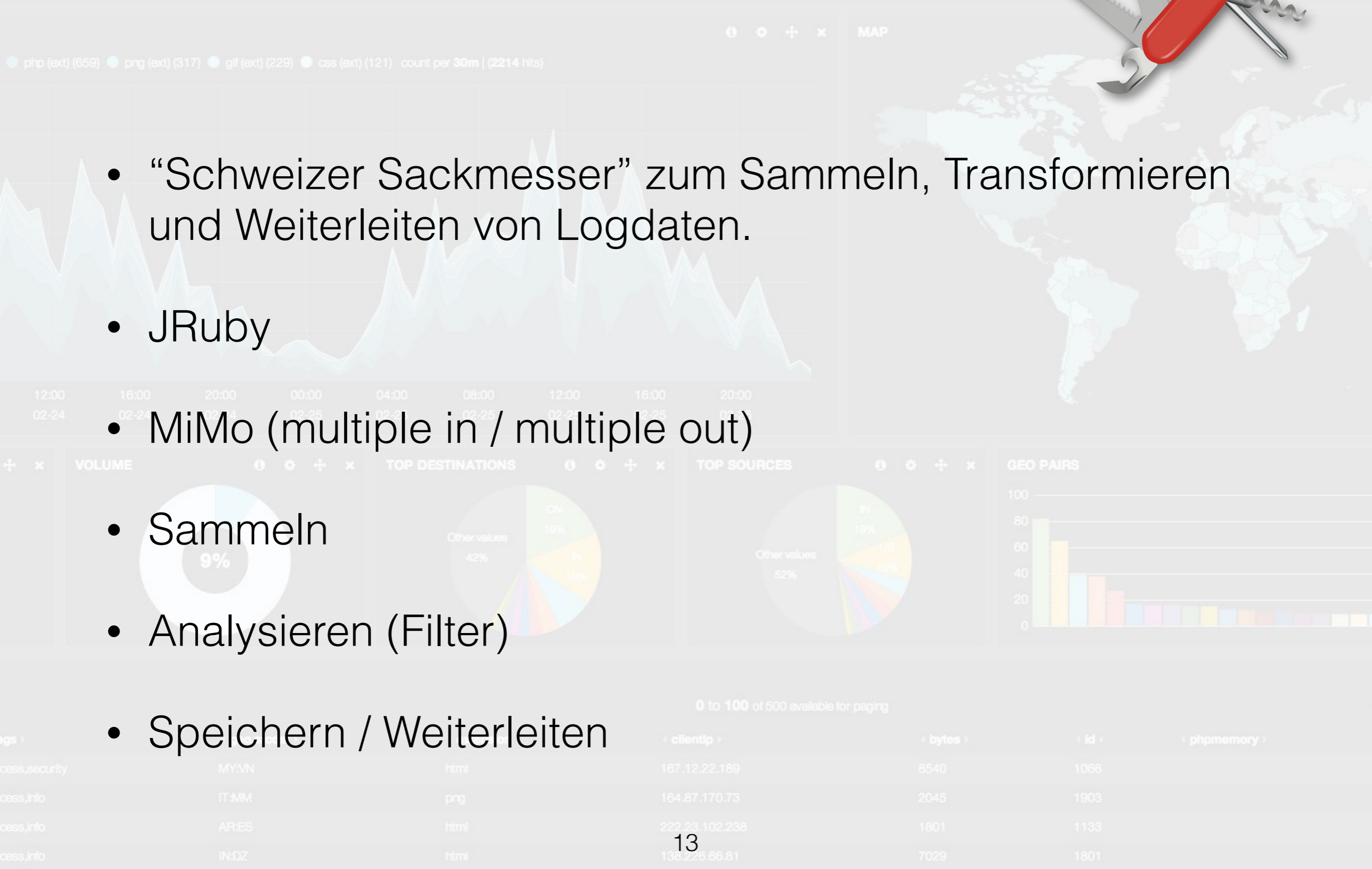
• JRuby

• MiMo (multiple in / multiple out)

• Sammeln

• Analysieren (Filter)

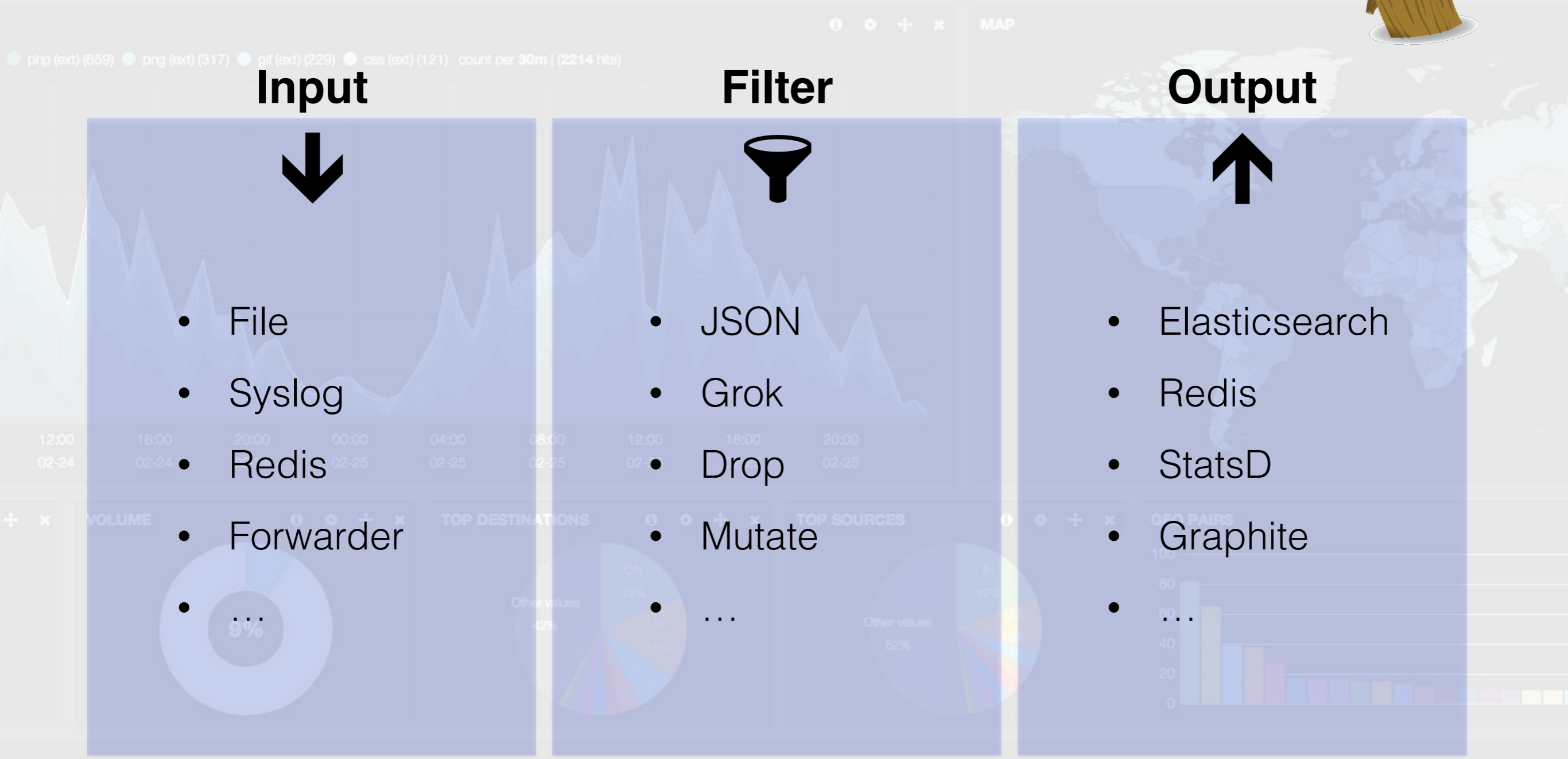
• Speichern / Weiterleiten





Logstash

SUCCESS



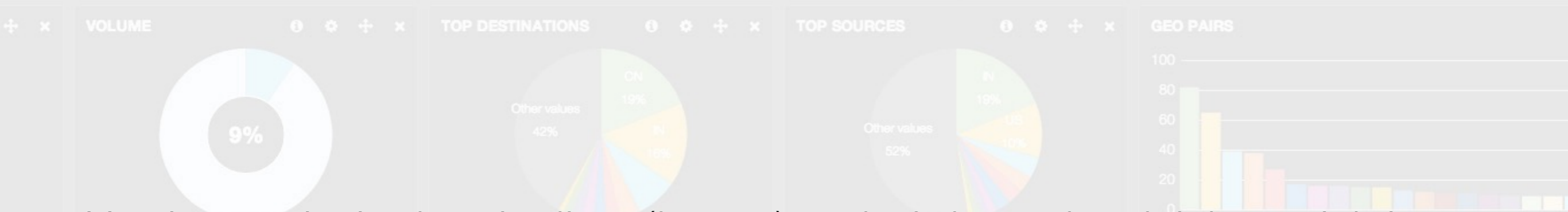
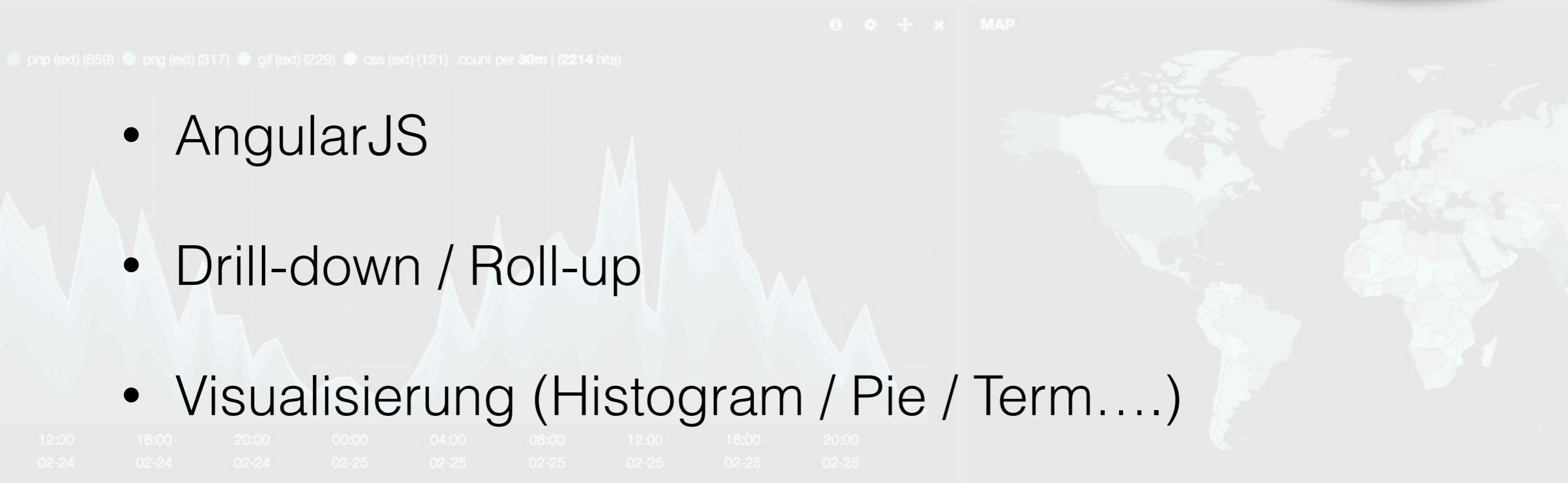
0 to 100 of 500 available for paging

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
cess,security	MY:VN	html	167.12.22.189	8540	1068	
cess,info	IT:MM	png	164.87.170.73	2045	1903	
cess,info	AR:ES	html	222.23.102.238	1801	1133	
cess,info	IN:DZ	html	138.226.66.81	7029	1801	



Kibana

- AngularJS
- Drill-down / Roll-up
- Visualisierung (Histogramm / Pie / Term....)



Version 4.0 ist in der pipeline, (beta 3) und wird aus dem leichtgewichtigen Javascript Frontend einen Server Prozess machen welcher dann aber auch (dank Elasticsearch 1.4) Daten aggregieren kann.

bytes	id	phpmemory
167.12.22.189	1068	
164.87.170.73	1903	
222.23.102.238	1133	
138.226.66.81	1801	



Kibana

SUCCESS

php (ext) (659) png (ext) (317) gif (ext) (229) css (ext) (121) count per 30m | (2214 hits)

Logstash Search 2 days ago to 3 minutes ago

QUERY **bytes:[0 TO 4000000] AND @tags:success**

FILTERING

EVENTS OVER TIME

View | Zoom Out | html (ext) (888) php (ext) (659) png (ext) (317) gif (ext) (229) css (ext) (121) count per 30m | (2214 hits)

MAP

REVENUE

22%

VOLUME

9%

TOP DESTINATIONS

CN 19%
IN 16%
Other values 42%

TOP SOURCES

IN 19%
US 10%
Other values 52%

GEO PAIRS

ALL EVENTS 0 to 100 of 500 available for paging

Fields	@tags	geo.srcdst	extension	clientip	bytes	id	phpmemory	response
success,security	MY:VN	html	167.12.22.189	8540	1066	200		
success,info	IT:MM	png	164.87.170.73	2045	1903	200		
success,info	AR:ES	html	222.23.102.238	1801	1133	200		
success,info	IN:DZ	html	138.226.66.81	7029	1801	200		

Tools



ElasticHQ

Cluster Overview

Cluster Statistics

- 6 Nodes
- 662 Total Shards
- 662 Successful Shards
- 25 Indices
- 766,106,100 Documents
- 469.1GB Size

Cluster Health

Status	Green
Timed Out?	false
# Nodes	6
# Data Nodes	3
Active Primary Shards	221
Active Shards	662
Relocating Shards	0
Initializing Shards	0
Unassigned Shards	0

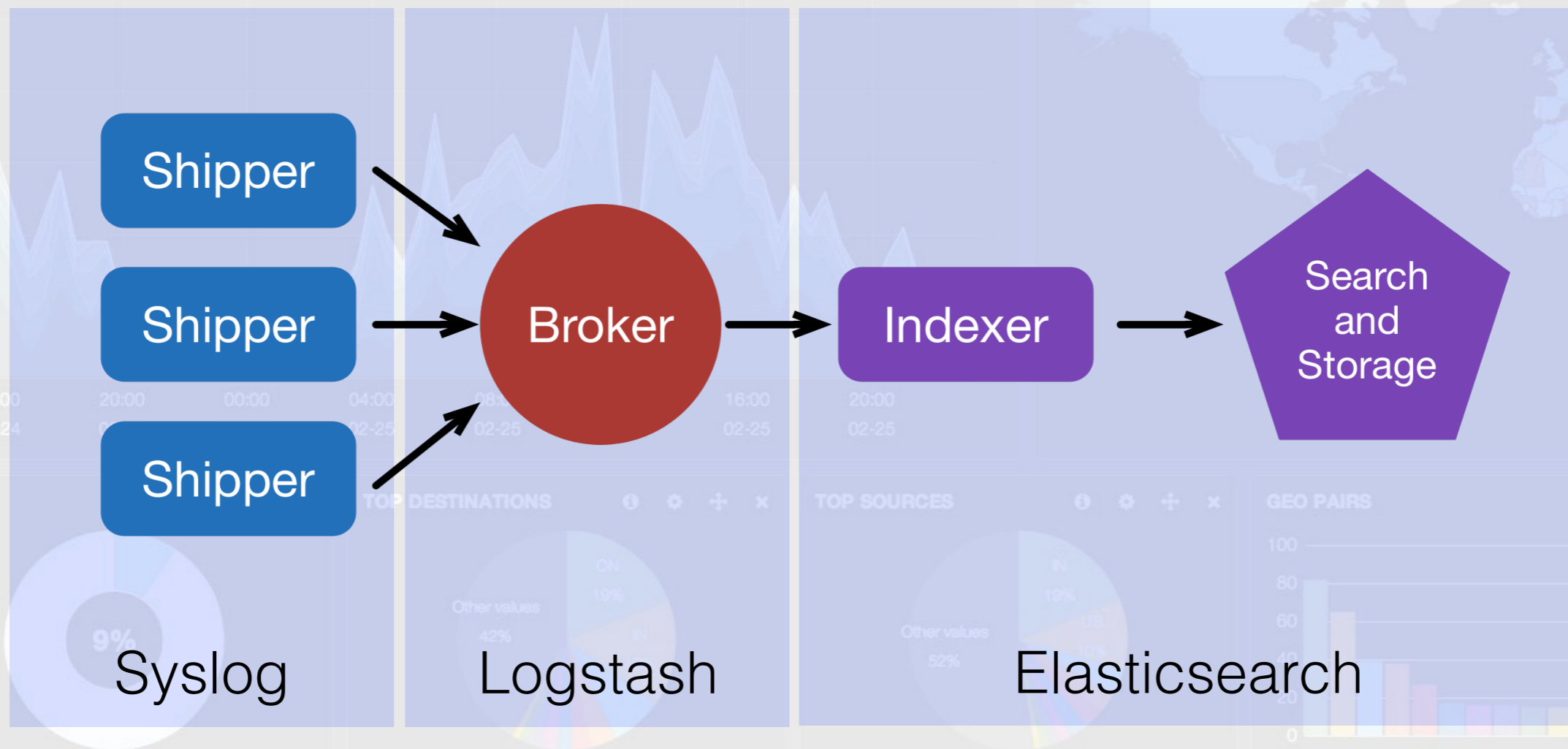
Indices

Index	# Docs	Primary Size	# Shards	# Replicas	Status
logstash-prd-2015.01.03	27,763,709	17.2GB	10	2	open
logstash-prd-2015.01.02	41,266,154	24.2GB	10	2	open
logstash-prd-2015.01.01	18,221,424	9.8GB	10	2	open
logstash-prd-2014.12.52	32,555,422	17.2GB	10	2	open
logstash-prd-2014.12.51	46,395,069	26.5GB	10	2	open
logstash-prd-2014.12.50	40,892,592	23.2GB	10	2	open
logstash-prd-2014.12.49	43,914,373	25.0GB	10	2	open
logstash-prd-2014.12.01	16,016,978	8.8GB	10	2	open
logstash-prd-2014.11.48	52,695,160	31.3GB	10	2	open
logstash-prd-2014.11.47	44,246,924	19.8GB	10	2	open

Curator

```
curator -host localhost -port 9200 delete --time-unit months --older-than 2
```

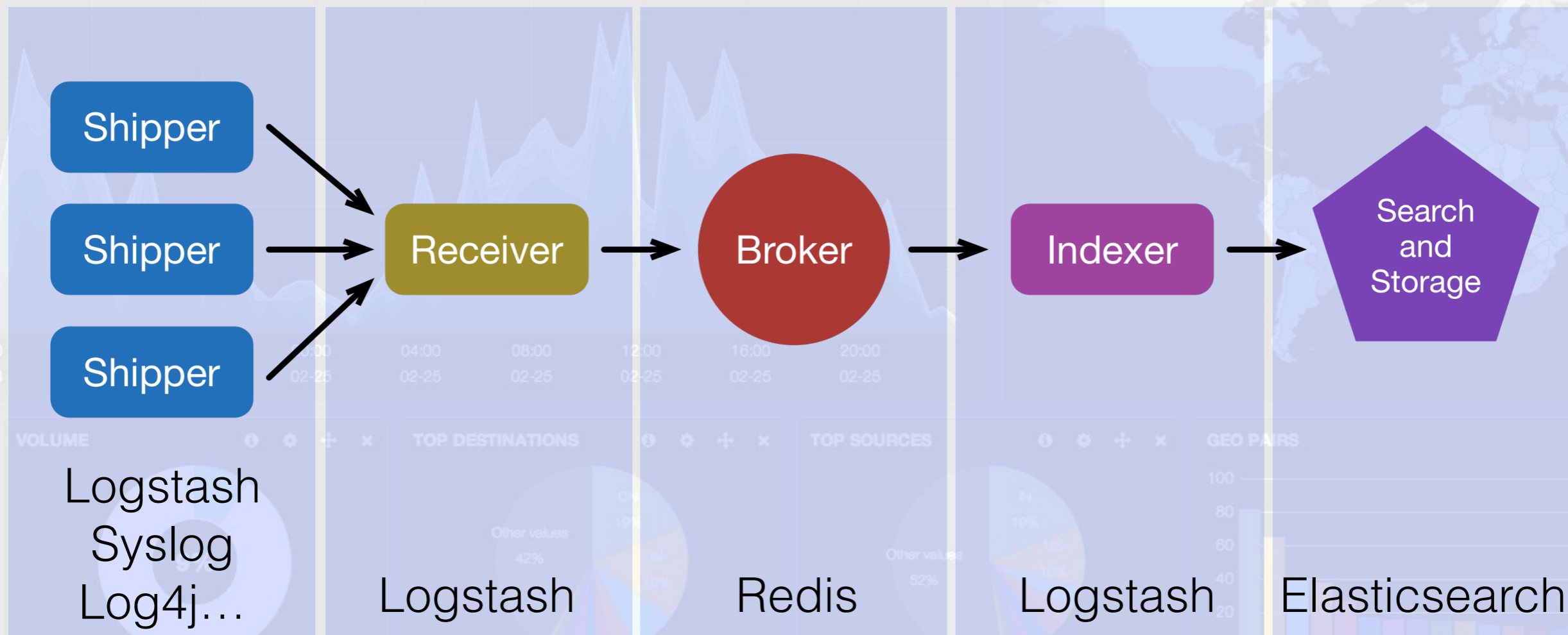
Architektur



0 to 100 of 500 available for paging

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
access_security	MY:VN	html	167.12.22.189	8540	1068	
access_info	IT:MM	png	164.87.170.73	2045	1903	
access_info	AR:ES	html	222.23.102.238	1801	1133	
access_info	IN:DZ	html	138.226.66.81	7029	1801	

Realistische Architektur :-)



SUCCESS

● php (ext) (659) ● png (ext) (317) ● gif (ext) (229) ● css (ext) (121) count per 30m | (2214 hits)

MAP

12:00
02-24

VOLUME TOP DESTINATIONS TOP SOURCES GEO PAIRS

Logstash
Syslog
Log4j...

Logstash

Redis

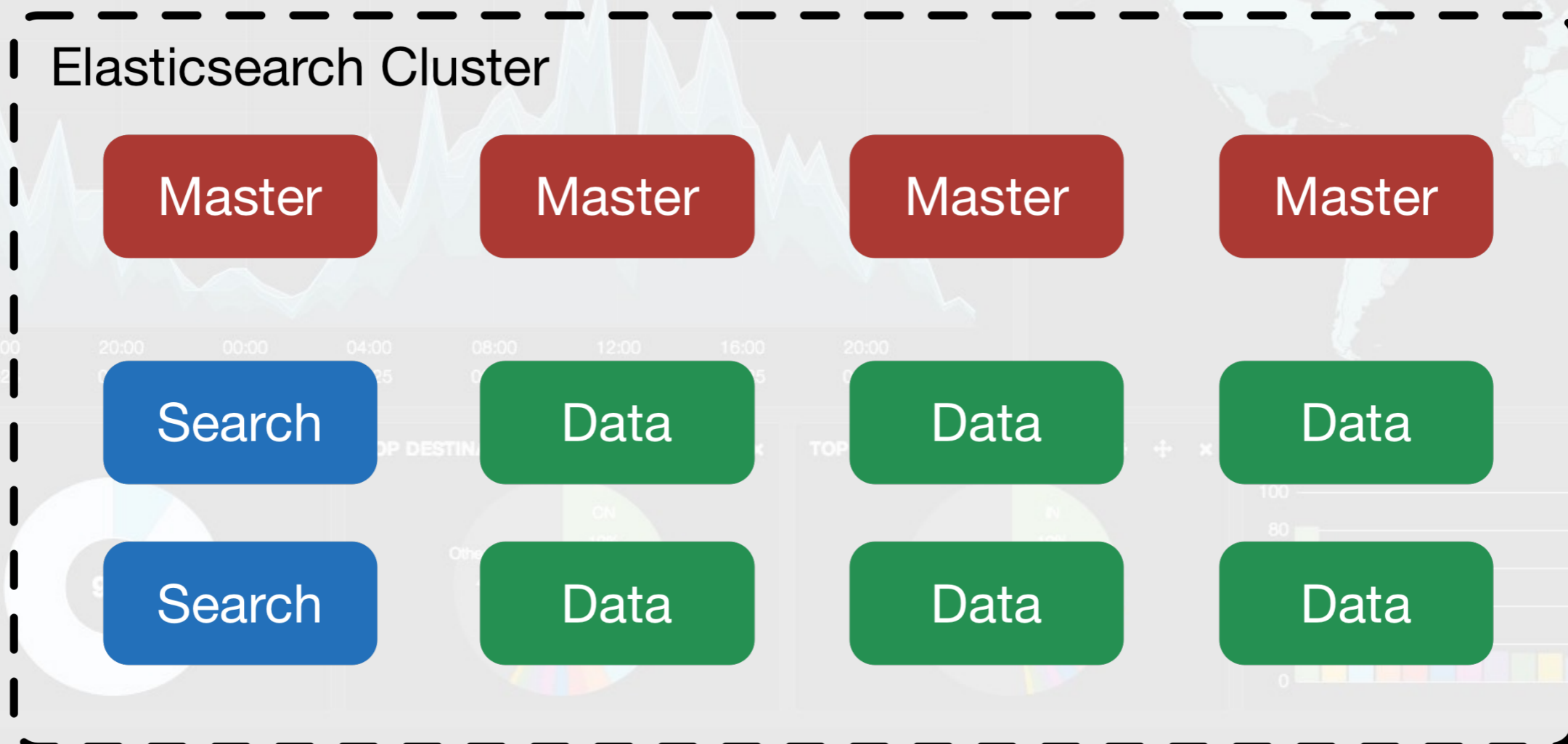
Logstash

Elasticsearch

0 to 100 of 500 available for paging

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
access_security	MY:VN	html	167.12.22.189	8540	1068	
access_info	IT:MM	png	164.87.170.73	2045	1903	
access_info	AR:ES	html	222.23.102.238	1801	1133	
access_info	IN:DZ	html	138.226.66.81	7029	1801	

Elasticsearch Node Architektur



SUCCESS

● php (ext) (659) ● png (ext) (317) ● gif (ext) (229) ● css (ext) (121) count per 30m | (2214 hits)

12:00 16:00 20:00 00:00 04:00 08:00 12:00 16:00 20:00
02-24 02-25

VOLUME

logs

access,security

access,info

access,info

access,info

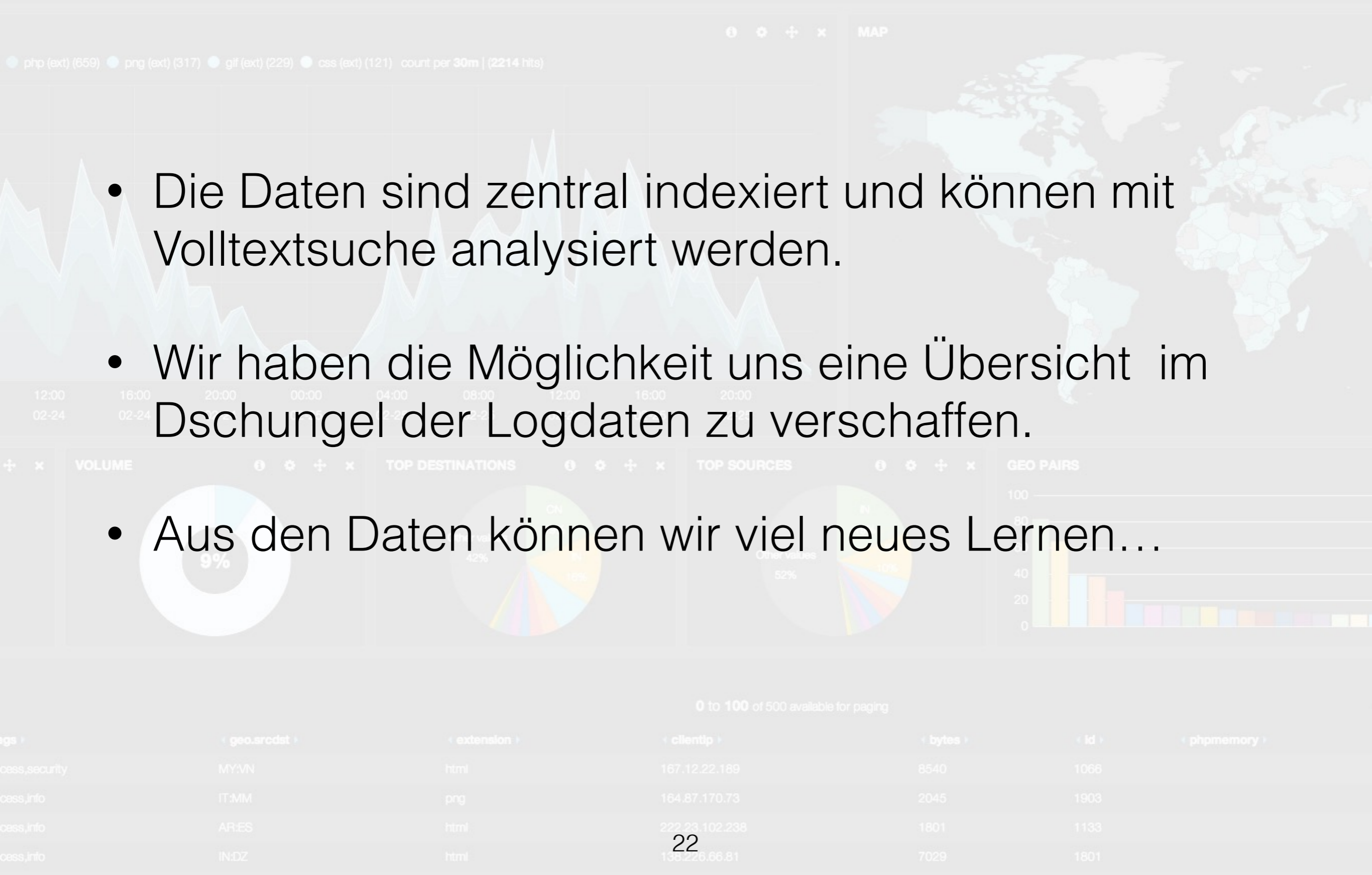
geo.srcdst	extension	clientip	bytes	id	phpmemory
MY:VN	html	167.12.22.189	8540	1068	
IT:MM	png	164.87.170.73	2045	1903	
AR:ES	html	222.23.102.238	1801	1133	
IN:IZ	html	135.226.66.81	7029	1801	

Übliche Fallstricke

- Nicht genügend Arbeitsspeicher (RAM, RAM, RAM...)
- Elasticsearch Server nutzt Swap Speicher!
- Langsamer Physischer Speicher.
- Java Memory Fragmentierung (Heap > 20GB)
- Elasticsearch mixed nodes (z.B.: master und data)
- Index Frequenz (Tag, Woche, Monat, Jahr)
- Shards und Copies (Split Brain)
- Logstash: CPU intensive Grok (regex) Ausdrücke.

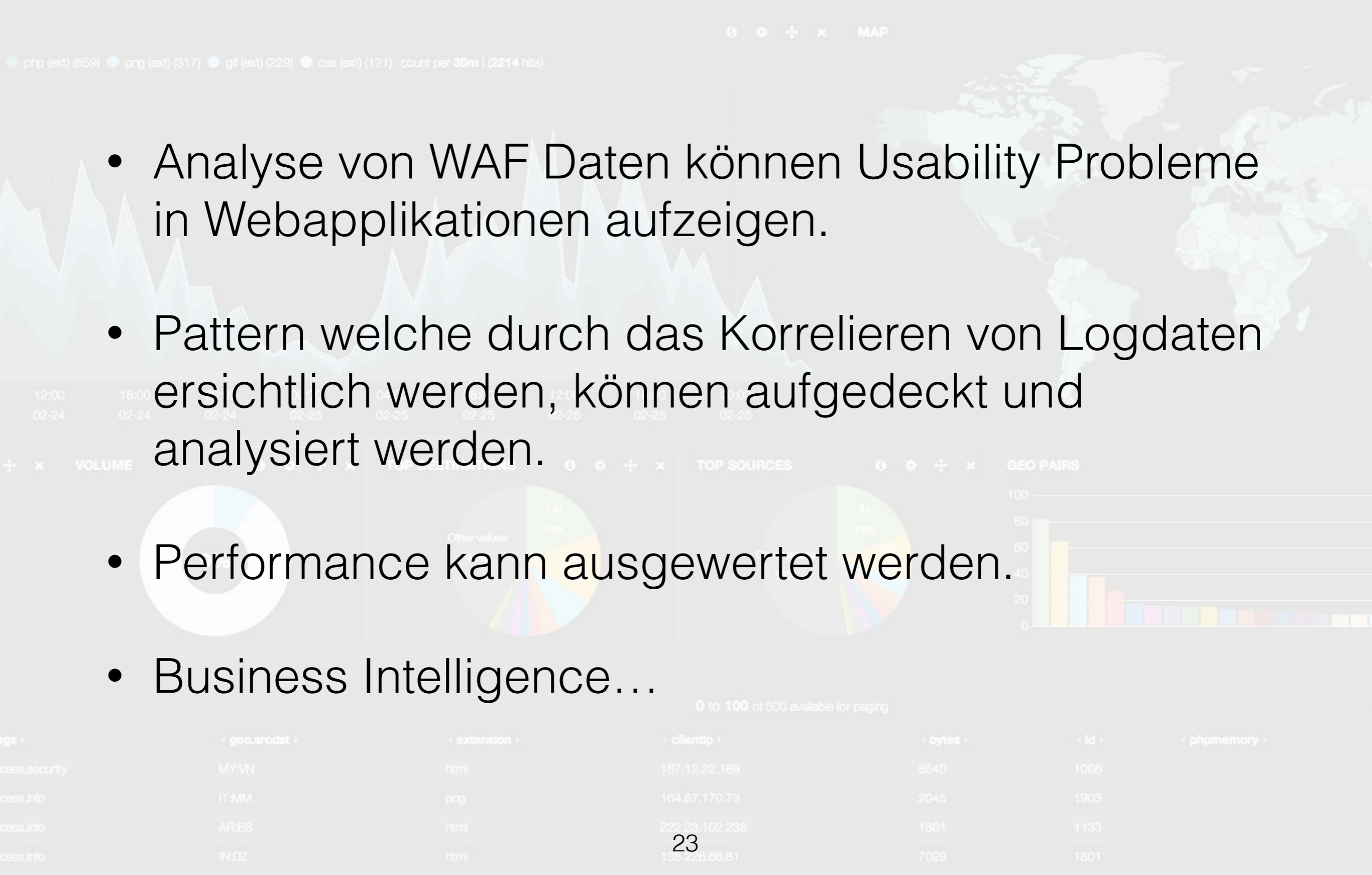
Und nun...

- Die Daten sind zentral indexiert und können mit Volltextsuche analysiert werden.
- Wir haben die Möglichkeit uns eine Übersicht im Dschungel der Logdaten zu verschaffen.
- Aus den Daten können wir viel neues Lernen...



Anwendung

- Analyse von WAF Daten können Usability Probleme in Webapplikationen aufzeigen.
- Pattern welche durch das Korrelieren von Logdaten ersichtlich werden, können aufgedeckt und analysiert werden.
- Performance kann ausgewertet werden.
- Business Intelligence...

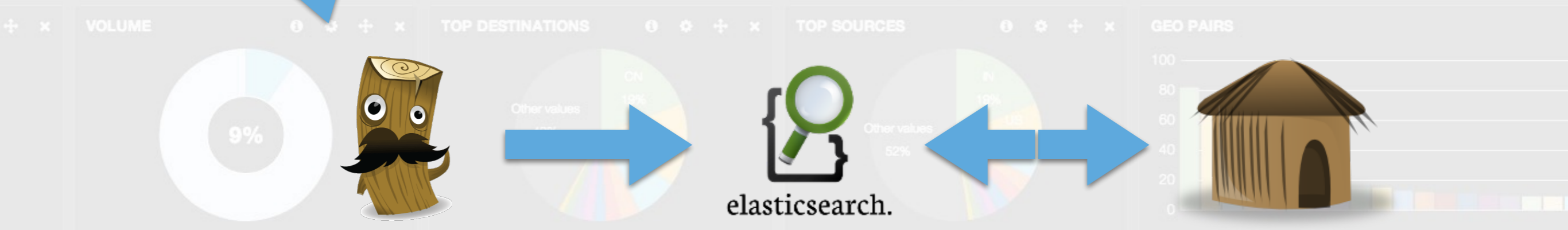
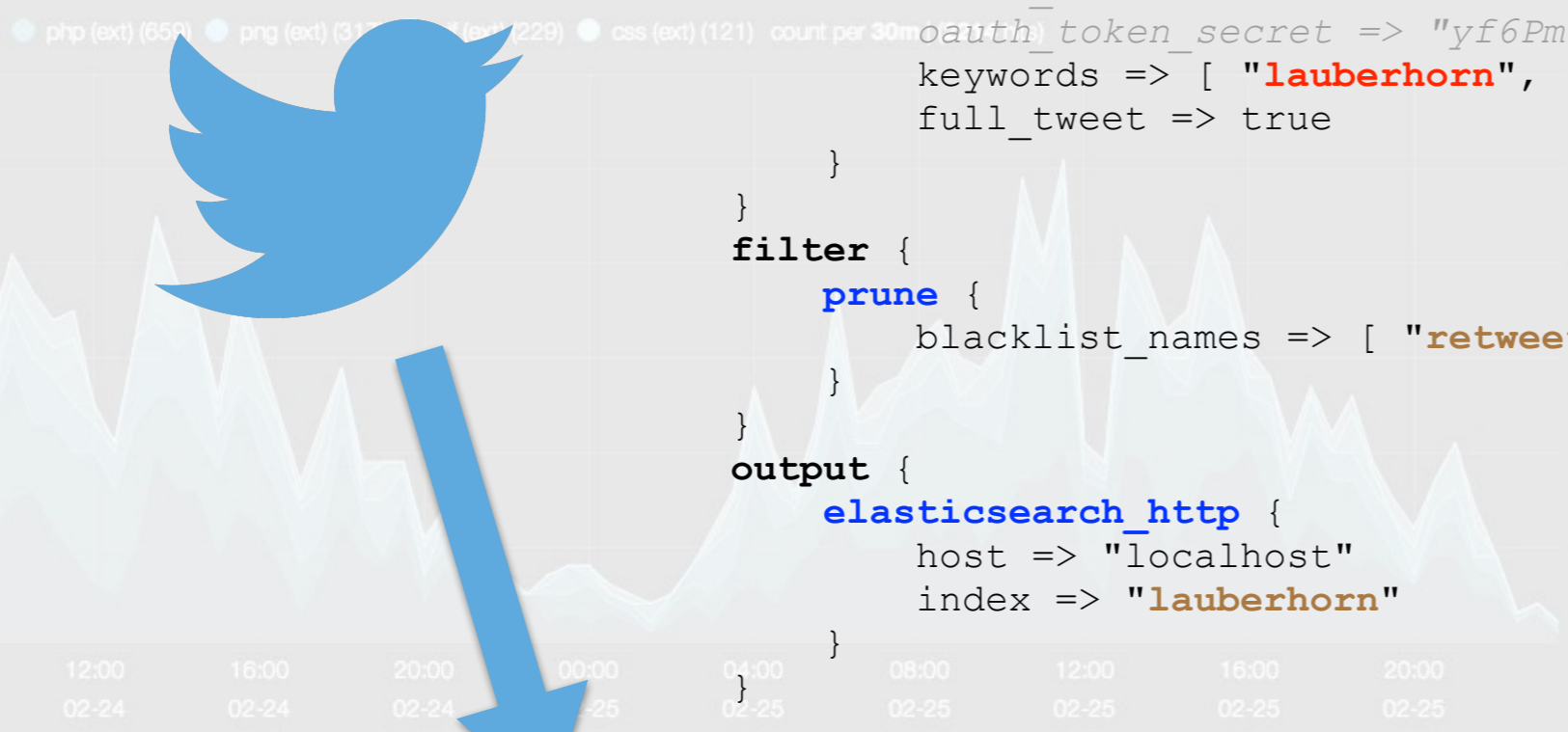


Demo

```

input {
  twitter {
    consumer_key => "vhFUbj4vccROMirZesU9Br"
    consumer_secret => "v9xeZ0tCggwo0zqJWDuR02KhWnh1W2SLgPkQJmKGYF2GRzw"
    oauth_token => "51665235-TASGSSMzAFyjGlmeHd0mmkYrqJuUw50TZImZf"
    oauth_token_secret => "yf6PmWqOZJMwdvsU0xOd3zRNvIWpOqxWAmcaSe47R"
    keywords => [ "lauberhorn", "wengen", "abfahrt", "skirennen" ]
    full_tweet => true
  }
}
filter {
  prune {
    blacklist_names => [ "retweeted_.*" ]
  }
}
output {
  elasticsearch_http {
    host => "localhost"
    index => "lauberhorn"
  }
}

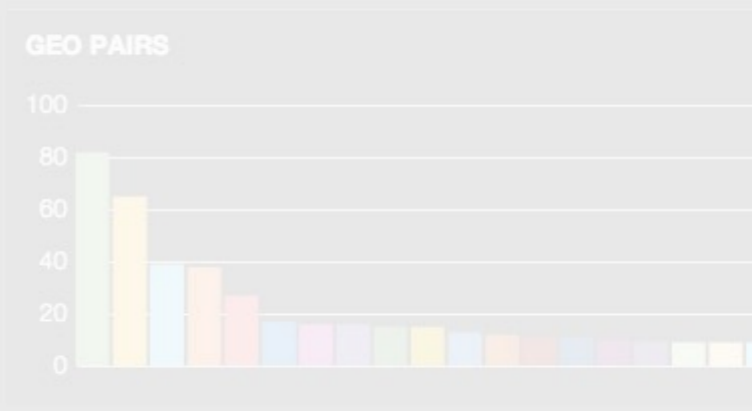
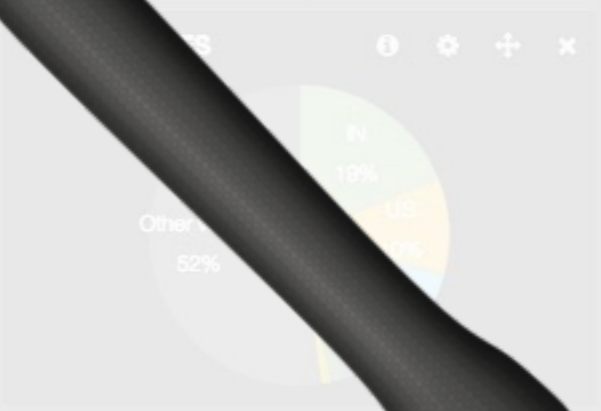
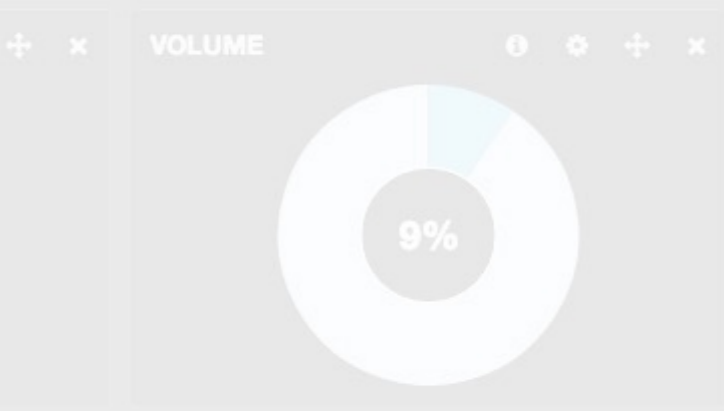
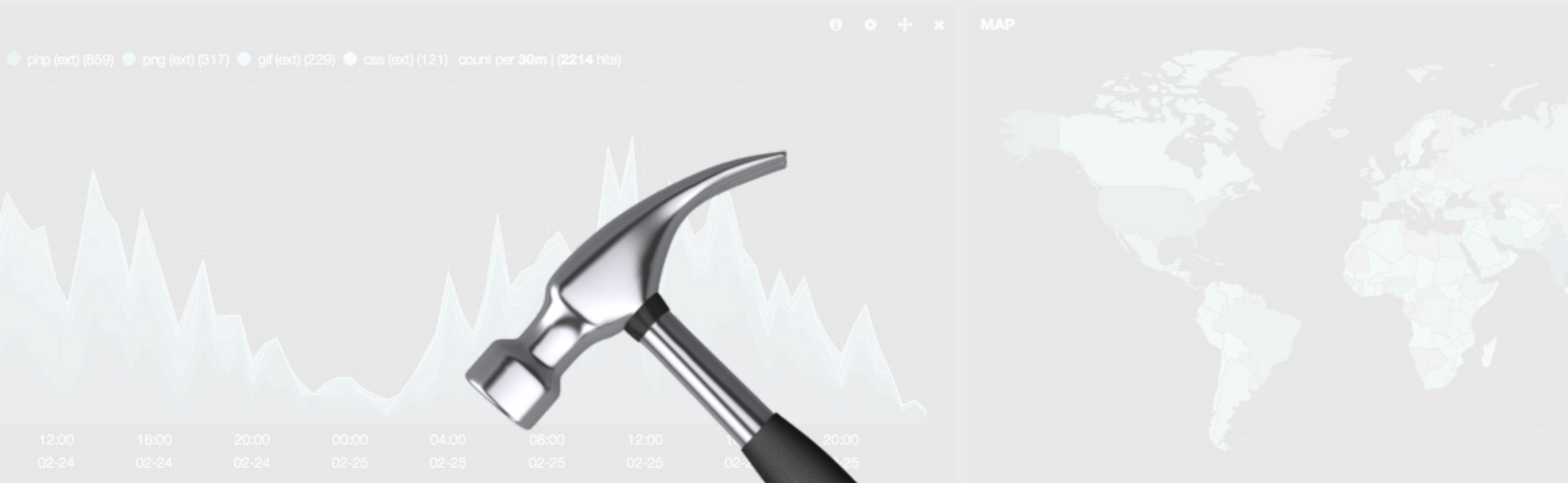
```



0 to 100 of 500 available for paging

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
access,security	MY:VN	html	167.12.22.189	8540	1068	
access,info	IT:MM	png	164.87.170.73	2045	1903	
access,info	AR:ES	html	222.23.102.238	1801	1133	
access,info	IN:DZ	html	135.226.66.81	7029	1801	

a fool with a tool is still a fool...



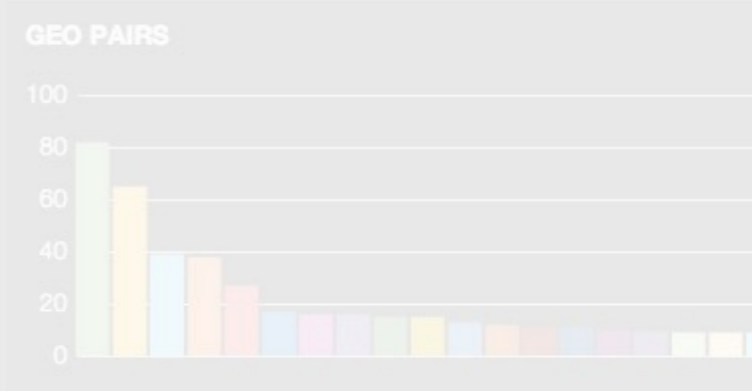
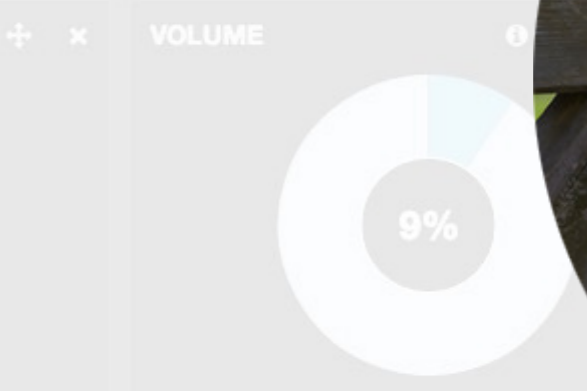
0 to 100 of 500 available for paging

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
access_security	MY:VN	html	167.12.22.189	8540	1068	
access_info	IT:MM	png	164.87.170.73	2045	1903	
access_info	AR:ES	html	222.23.102.238	1801	1133	
access_info	IN:DZ	html	135.226.66.81	7029	1801	

SUCCESS

php (ext) (659) png (ext) (317) gif (ext) (229) css (ext) (121) count per 30m | (2214 hits)

Any Questions ?



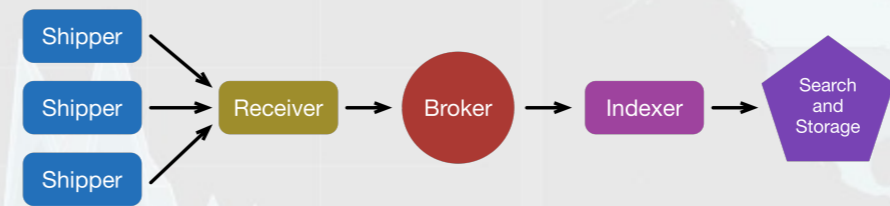
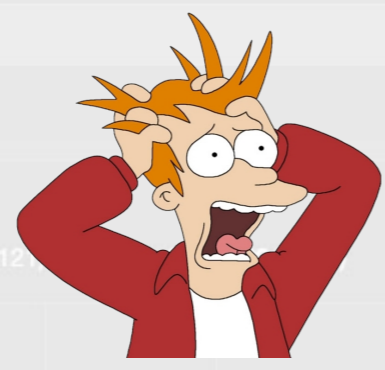
100 of 500 available for paging

logs	geo.srcdst	extension	clientip	bytes	id	phpmemory
access_security	MY:VN	html	167.12.22.189	8540	1068	
access_info	IT:MM	png	164.87.170.73	2045	1903	
access_info	AR:ES	html	222.23.102.238	1801	1133	
access_info	IN:DZ	html	135.226.66.81	7029	1801	

SUCCESS



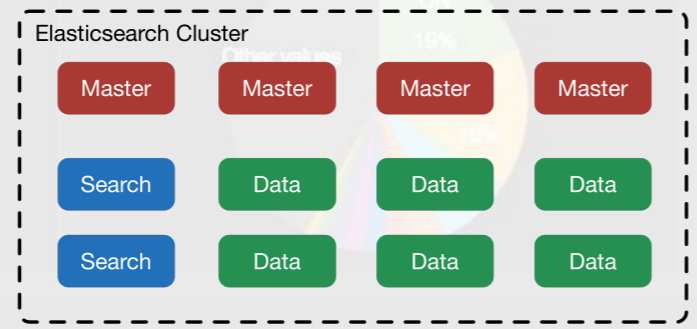
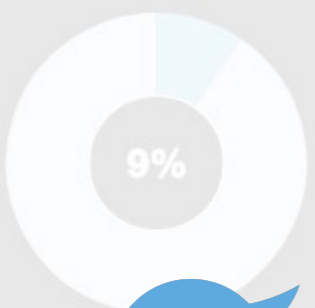
elasticsearch.



DANKE



VOLUME TOP DESTINATIONS TOP SOURCES GEO PAIRS



extension	clientip
html	167.12.22.189
png	164.87.170.73
html	222.23.102.238
html	138.226.66.81