



様々なメトリクスやログを集めてシステム解析 - Elastic Stackの入門と活用 -

2018/10/27

Community Engineer @Elastic
Jun Ohtani @johtani

アジェンダ

- メトリック／ログとは？
- システムメトリクス解析、ログ解析を試しにやってみよう
 - Beats - Elasticsearch - Kibanaで解析
- 本格的に解析をやるには？
 - Logstashでログやメトリクスを中継・集約
- さらに色々試してみるには？

about

- Me, Jun Ohtani / Community Engineer
 - lucene-gosenコミッター
 - データ分析基盤構築入門 共著
 - <http://blog.johtani.info>



- Elastic, founded in 2012
 - Products: Elasticsearch, Logstash, Kibana, Beats
Elastic APM,
Elastic Cloud, Swiftype
 - Professional services: Support & development subscriptions
Trainings, Consulting, SaaS





どんなメトリック、
ログを集めていますか？

メトリック

- CPU、メモリ使用率、ディスク使用率
- アクセス数、ネットワーク転送量
- 応答時間
- コネクション数
- トランザクション数、売上
- コンテナの上の各種メトリクス

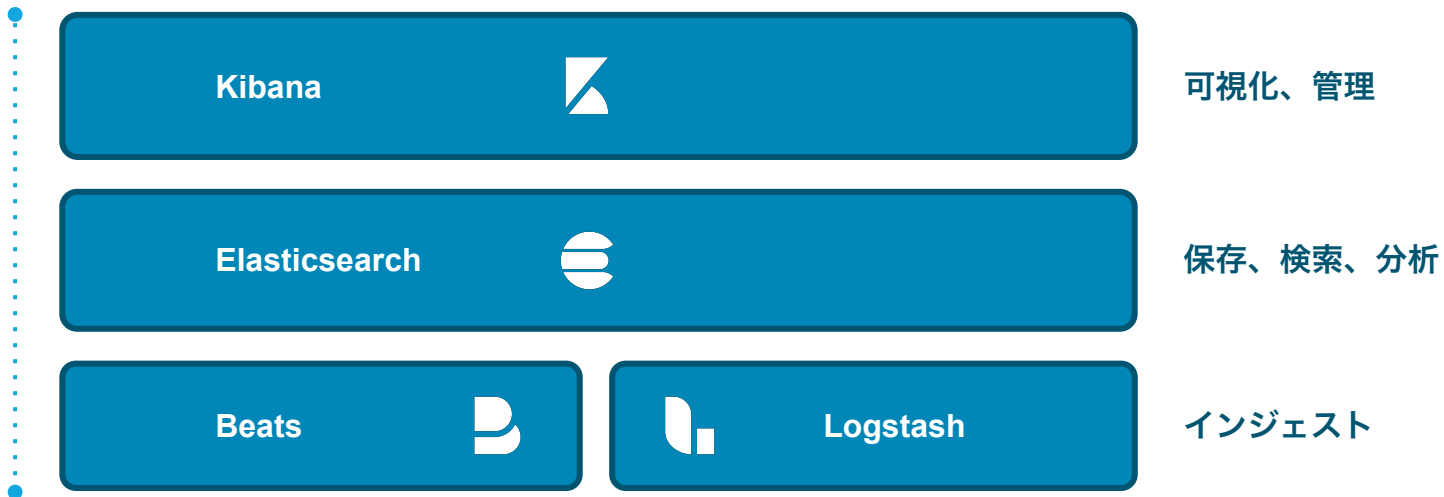
ログ

- 認証ログ
- システムログ
- アプリケーションログ
- Slow log
- アクセスログ
- コンテナの中のログ



できればログとメトリックを
まとめて1つの画面で
見たいですね？

Elastic Stack





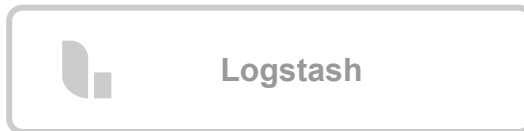
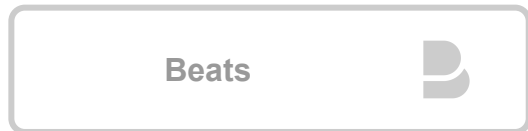
ソリューション



可視化、管理



保存、検索、分析



インジェスト



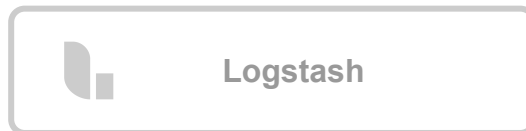
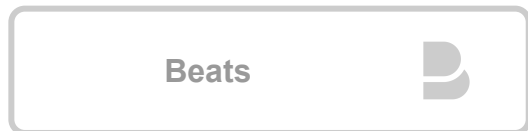
ソリューション



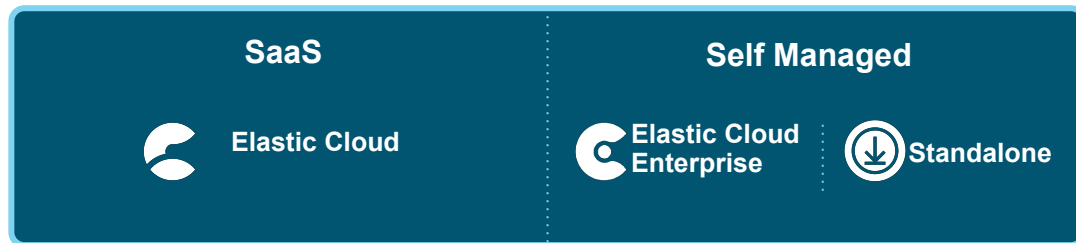
可視化、管理



保存、検索、分析



インジェスト



デプロイ



ソリューション



可視化、管理



保存、検索、分析



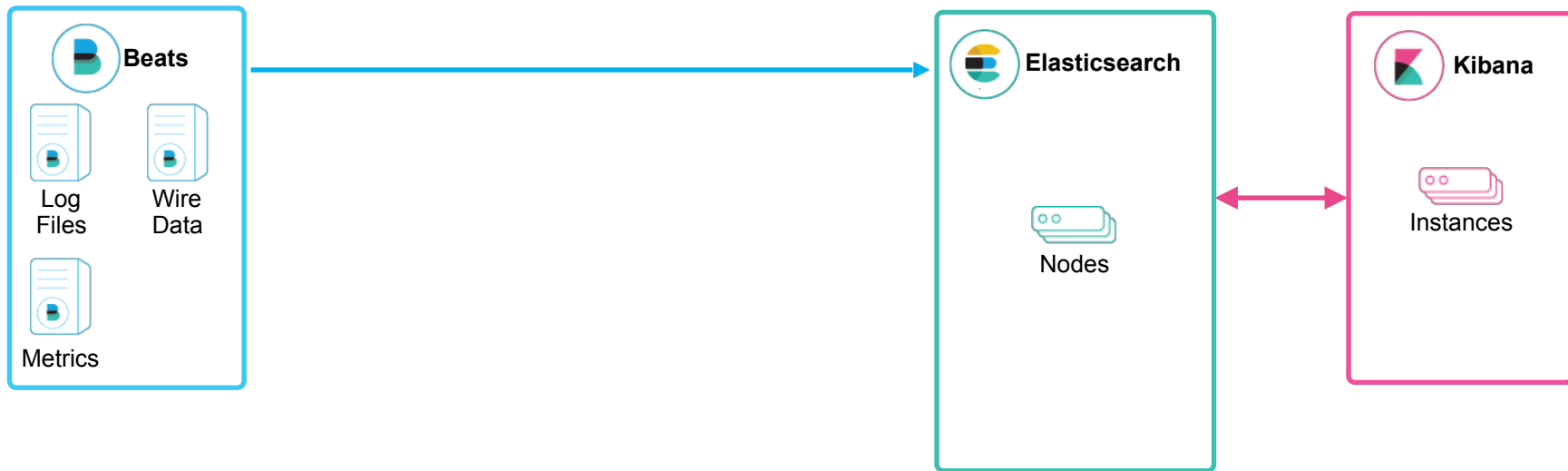
インジェスト



デプロイ

定型のメトリクス/ログ解析を Elastic Stackで

メトリック・ログ分析（簡易版）



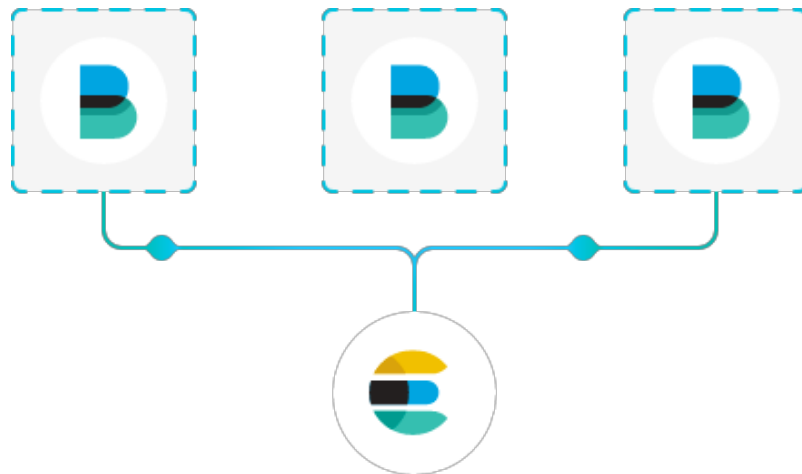


beats



Beats

軽量データシッパー



ソースからデータを転送

転送しElasticsearchに集約

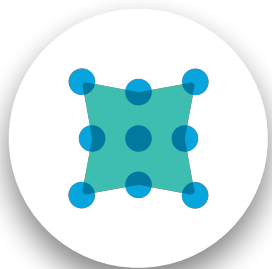
変換とパースのため
Logstashに転送

Elastic Cloudに転送

Libbeat: カスタムbeatsのた
めのAPIフレームワーク

30以上のコミュニティbeats

The Beats family



Packetbeat

Network data



Metricbeat

Metrics



Winlogbeat

Windows Event Logs



Auditbeat

Audit data



Filebeat

Log files



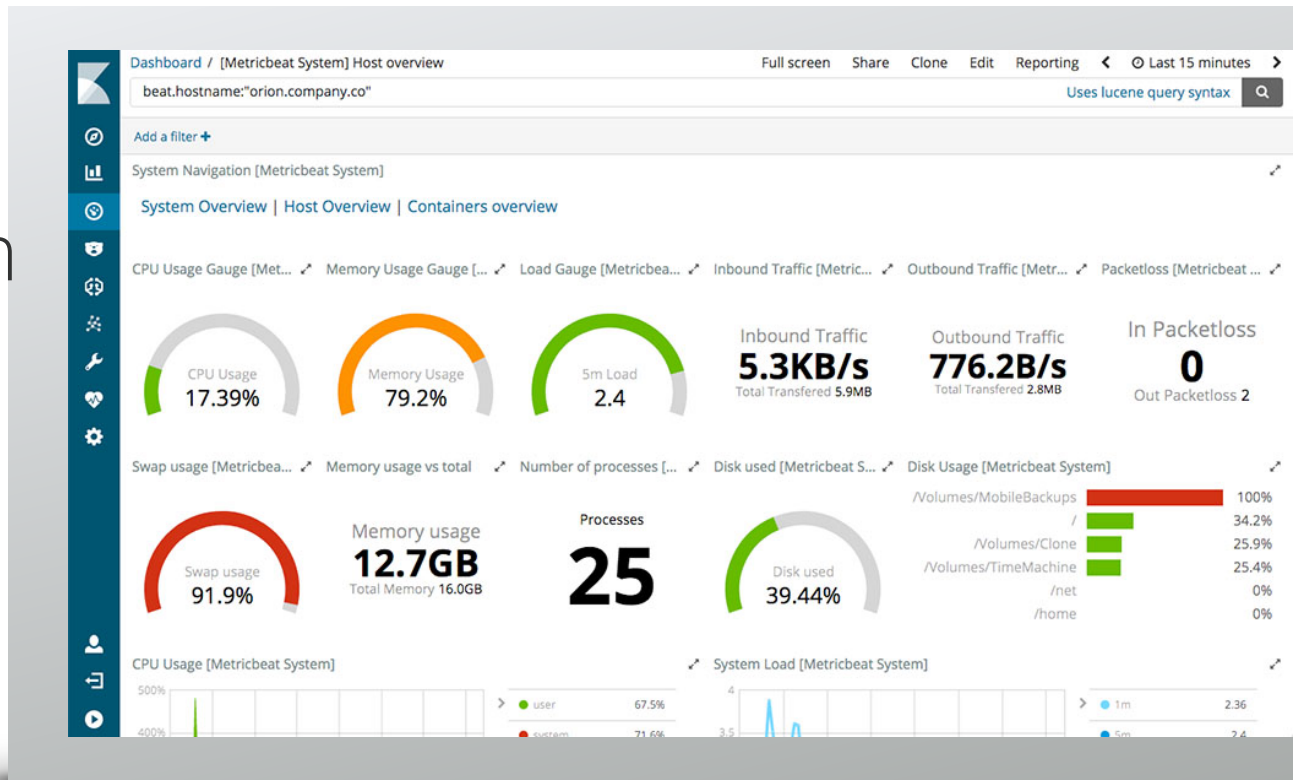
Heartbeat

Uptime monitoring

+40
community
Beats

Metricbeat

Collect system
and application
metrics



Metricbeat

lots of **modules**



System



Apache



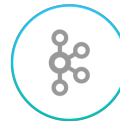
Docker



NGINX



HAProxy



Kafka



MongoDB



MySQL



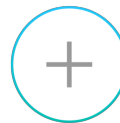
PostgreSQL



Prometheus



Jolokia



Add your own

Metricbeat モジュール

[Aerospike module](#)

[Apache module](#)

[Ceph module](#)

[Couchbase module](#)

[Docker module](#)

[Dropwizard module](#)

[Elasticsearch module](#)

[Etc module](#)

[Golang module](#)

[Graphite module](#)

[HAProxy module](#)

[HTTP module](#)

[Jolokia module](#)

[Kafka module](#)

[Kibana module](#)

[Kubernetes module](#)

[kvm module](#)

[Logstash module](#)

[Memcached module](#)

[MongoDB module](#)

[Munin module](#)

[MySQL module](#)

[Nginx module](#)

[PHP_FPM module](#)

[PostgreSQL module](#)

[Prometheus module](#)

[RabbitMQ module](#)

[Redis module](#)

[System module](#)

[uwsgi module](#)

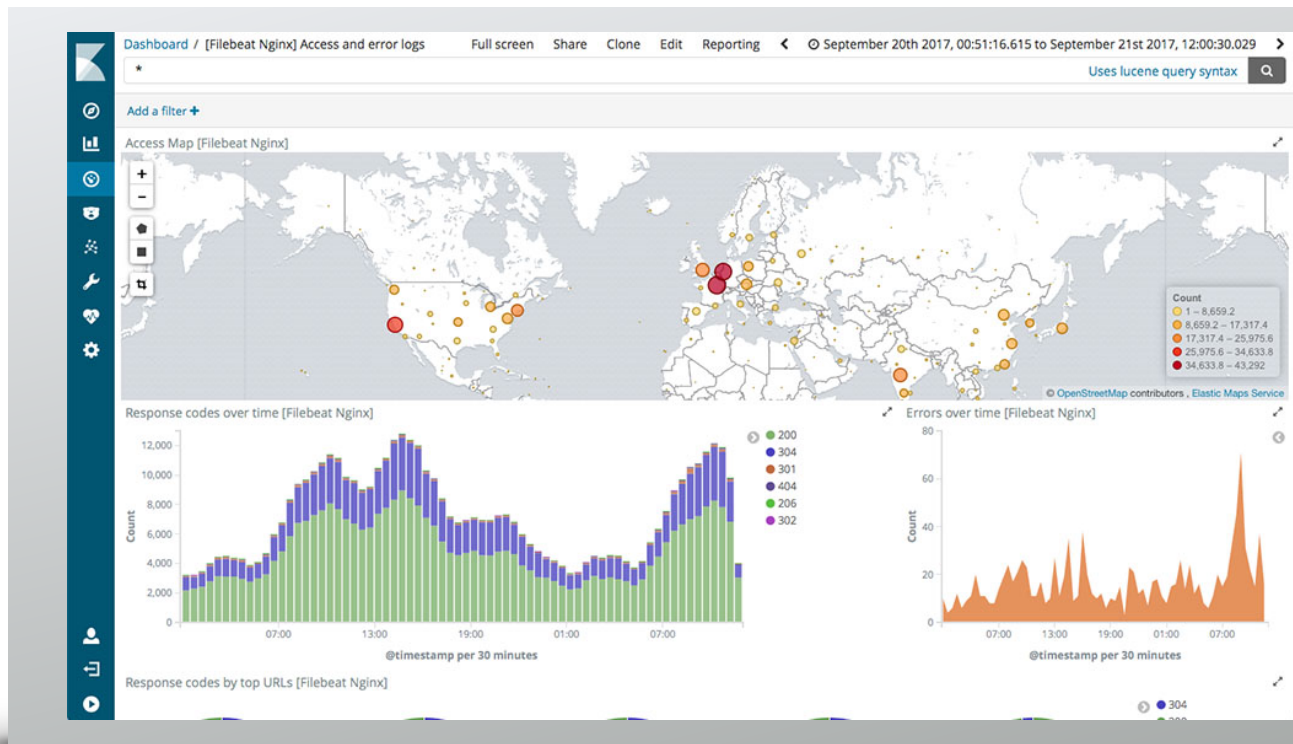
[vSphere module](#)

[Windows module](#)

[ZooKeeper module](#)

Filebeat

tail log from
file



many **modules**



Apache



Nginx



Auditd



MySQL

Filebeat modules - v6.4.2

[Apache2 module](#)

[Auditd module](#)

[Icinga module](#)

[IIS module](#)

[Kafka module](#)

[Logstash module](#)

[MongoDB module](#)

[MySQL module](#)

[Nginx module](#)

[Osquery module](#)

[PostgreSQL module](#)

[Redis module](#)

[System module](#)

[Traefik module](#)

Packetbeat

Capture the Packet

```
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on en0, link-type EN10MB (Ethernet), capture size 65535 bytes
10:03:59.594512 IP 172.31.98.131.65048 > nuq04s19-in-f21.1e100.net.https: UDP, length 24
10:03:59.692308 IP nuq04s19-in-f21.1e100.net.https > 172.31.98.131.65048: UDP, length 36
10:03:59.726313 IP 172.31.98.131.60568 > r-199-59-148-82.twtrr.com.https: Flags [..], ack 1987817713, win 4096, length 0
10:03:59.801353 IP r-199-59-148-82.twtrr.com.https > 172.31.98.131.60568: Flags [..], ack 1, win 1456, options [nop,nop,TS val 1737158165 ecr 1065051819], length 0
10:03:59.912168 IP pc-in-f189.1e100.net.https > 172.31.98.131.60078: Flags [P..], seq 391100909:391100994, ack 1961900067, win 1651, options [nop,nop,TS val 182273890 ecr 1065405533], length 85
10:03:59.912231 IP 172.31.98.131.60078 > pc-in-f189.1e100.net.https: Flags [..], ack 85, win 4093, options [nop,nop,TS val 1065411882 ecr 182273890], length 0
10:04:00.383581 IP 172.31.98.131.57399 > google-public-dns-a.google.com.domain: 48543; PTR: 131.98.31.172.in-addr.arpa. (44)
10:04:00.466579 IP google-public-dns-a.google.com.domain > 172.31.98.131.57399: 48543 NXDomain 0/0/0 (44)
10:04:00.467926 IP 172.31.98.131.52072 > google-public-dns-a.google.com.domain: 9347; PTR: 53.239.125.74.in-addr.arpa. (44)
10:04:00.568618 IP google-public-dns-a.google.com.domain > 172.31.98.131.52072: 9347 1/0/0 PTR nuq04s19-in-f21.1e100.net. (83)
10:04:00.569672 IP 172.31.98.131.59451 > google-public-dns-a.google.com.domain: 6362; PTR: 82.148.59.109.in-addr.arpa. (44)
10:04:00.676626 IP google-public-dns-a.google.com.domain > 172.31.98.131.59451: 6362 1/0/0 PTR r-199-59-148-82.twtrr.com. (83)
10:04:00.677667 IP 172.31.98.131.52322 > google-public-dns-a.google.com.domain: 26687; PTR: 189.28.125.74.in-addr.arpa. (44)
10:04:00.769797 IP google-public-dns-a.google.com.domain > 172.31.98.131.52322: 26687 1/0/0 PTR pc-in-f189.1e100.net. (78)
10:04:01.230731 IP 172.31.98.131.49573 > pb-in-f95.1e100.net.http: Flags [..], ack 3226625146, win 4096, length 0
10:04:01.348942 IP pb-in-f95.1e100.net.http > 172.31.98.131.49573: Flags [..], ack 1, win 341, options [nop,nop,TS val 4158964323 ecr 1065277921], length 0
10:04:01.367564 IP 172.31.98.131.59991 > pc-in-f125.1e100.net.jobber-client: Flags [P..], seq 53622692:53622809, ack 3725017102, win 65535, length 117
10:04:01.511834 IP pc-in-f125.1e100.net.jobber-client > 172.31.98.131.59991: Flags [P..], seq 1:134, ack 117, win 65100, length 133
10:04:01.511834 IP 172.31.98.131.59991 > pc-in-f125.1e100.net.jobber-client: Flags [..], ack 134, win 65535, length 0
10:04:01.778555 IP 172.31.98.131.49474 > google-public-dns-a.google.com.domain: 40324; PTR: 8.8.8.8.in-addr.arpa. (38)
10:04:01.871839 IP google-public-dns-a.google.com.domain > 172.31.98.131.49474: 40324 1/0/0 PTR google-public-dns-a.google.com. (82)
10:04:01.872628 IP 172.31.98.131.50753 > google-public-dns-a.google.com.domain: 14329; PTR: 95.79.194.173.in-addr.arpa. (44)
10:04:01.907102 IP 172.31.98.131.49578 > 199.27.19.134.http: Flags [..], ack 682580952, win 4096, length 0
```

File Edit View Go Capture Analyze Statistics Telephony Tools Internals Help

Filter: ip.addr == 192.168.1.6 Expression... Clear Apply

No.	Time	Source	Destination	Protocol	Info
19511	995.233558000	192.168.1.6	8.8.8.8	DNS	Standard query A download340.avast.com
19512	995.233597000	192.168.1.8	192.168.1.6	TCP	Redirect (Redirect for host)
19513	995.233631000	192.168.1.6	8.8.8.8	DNS	Standard query A download340.avast.com
19514	995.248689000	8.8.8.8	192.168.1.6	DNS	Standard query response A 82.192.95.92
19515	995.248710000	8.8.8.8	192.168.1.6	DNS	Standard query response A 82.192.95.92
19516	995.260447000	192.168.1.6	82.192.95.92	TCP	55552 > http [FIN, ACK] Seq=200 Ack=1154 Win=16368 Len=0
19520	995.312985000	82.192.95.92	192.168.1.6	TCP	http > 55555 [SYN, ACK] Seq=0 Ack=1 Win=5840 Len=0 HSS=1460 SACK PerM=1 WS=128
19521	995.313009000	82.192.95.92	192.168.1.6	TCP	http > 55555 [SYN, ACK] Seq=0 Ack=1 Win=5840 Len=0 HSS=1460 SACK PerM=1 WS=128
19522	995.313438000	192.168.1.6	82.192.95.92	TCP	55555 > http [ACK] Seq=1 Ack=1 Win=17520 Len=0
19523	995.314191000	192.168.1.6	82.192.95.92	TCP	[TCP Dup ACK 19522] 55555 > http [ACK] Seq=1 Ack=1 Win=17520 Len=0
19524	995.324631000	192.192.95.92	192.168.1.6	TCP	http > 55552 [ACK] Seq=1154 Ack=201 Win=6912 Len=0
19525	995.324668000	82.192.95.92	192.168.1.6	TCP	[TCP Dup ACK 19524] 1154 > 55552 [ACK] Seq=1154 Ack=201 Win=6912 Len=0
19527	995.325988000	192.168.1.6	82.192.95.92	TCP	[TCP segment of a reassembled PDU]
19528	995.326011000	192.168.1.6	82.192.95.92	TCP	[TCP segment of a reassembled PDU]
19529	995.326039000	192.168.1.6	82.192.95.92	HTTP	POST /cgi-bin/lav54stats.cgi HTTP/1.1 (lav54/stats)
19530	995.326278000	192.168.1.6	82.192.95.92	TCP	[TCP Retransmission] [TCP segment of a reassembled PDU]
19531	995.379511000	82.192.95.92	192.168.1.6	TCP	http > 55555 [ACK] Seq=1 Ack=206 Win=6912 Len=0
19532	995.379525000	82.192.95.92	192.168.1.6	TCP	http > 55555 [ACK] Seq=1 Ack=206 Win=6912 Len=0
19533	995.380658000	82.192.95.92	192.168.1.6	TCP	http > 55555 [ACK] Seq=1 Ack=1104 Win=8832 Len=0
19534	995.380712000	82.192.95.92	192.168.1.6	TCP	[TCP Dup ACK 19533] 1104 > 55555 [ACK] Seq=1 Ack=1104 Win=8832 Len=0
19535	995.382931000	82.192.95.92	192.168.1.6	HTTP	HTTP/1.1 204 No Content
19536	995.382911000	82.192.95.92	192.168.1.6	HTTP	[TCP Retransmission] HTTP/1.1 204 No Content
19539	995.505191000	192.168.1.6	82.192.95.92	TCP	55555 > http [RST, ACK] Seq=1104 Ack=93 Win=0 Len=0
19540	995.505211000	192.168.1.6	82.192.95.92	TCP	55555 > http [RST, ACK] Seq=1104 Ack=93 Win=0 Len=0
19559	996.308299000	192.168.1.6	149.7.96.236	TCP	55553 > mtp [SYN] Seq=0 Win=8192 Len=0 HSS=1460 WS=4 SACK PerM=1
19562	996.308324000	192.168.1.8	192.168.1.6	TCP	Redirect (Redirect for host)
19562	996.308393000	192.168.1.6	149.7.96.236	TCP	55553 > mtp [SYN] Seq=0 Win=8192 Len=0 HSS=1460 WS=4 SACK PerM=1

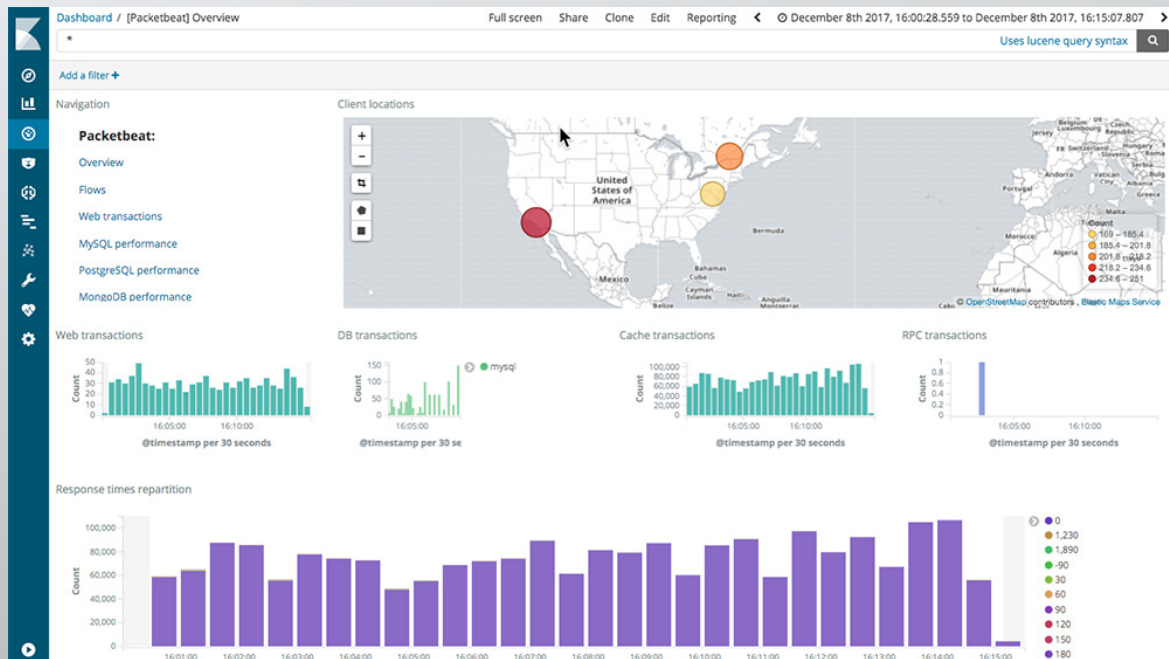
Frame 9164: 77 bytes on wire (616 bits), 77 bytes captured (616 bits) on 0
Ethernet II, Src: HonNaiPr, 26:b5:30 (c8:cb:38:26:b5:30), Dst: Azurewav, 43:90:de (00:15:af:43:90:de)
Internet Protocol Version 4, Src: 68.126.7.59 (68.126.7.59), Dst: 192.168.1.6 (192.168.1.6)
Transmission Control Protocol, Src Port: 19207 (19207), Dst Port: 55400 (55400), Seq: 1, Ack: 1, Len: 23

```
0000 00 15 af 43 90 de c8 38 26 b5 30 08 00 45 00 ...C.... 86.0.1.e.  
0010 00 37 57 57 40 00 ef 06 26 f0 44 7e 07 3b c0 a8 79m0... &D.;...  
0020 01 06 40 07 08 68 00 00 00 0f 49 3f 88 50 14 ...K..h...;...P.  
0030 00 00 5a 16 00 00 00 00 00 00 00 00 00 00 ...;...;...;...  
0040 65 27 72 65 20 6e 6f 74 20 68 6f 6d 65 e're not home
```

eth1: <live capture in progress> File: Packets: 19552 Displayed: 5155 Marked: 0 Profile: Default

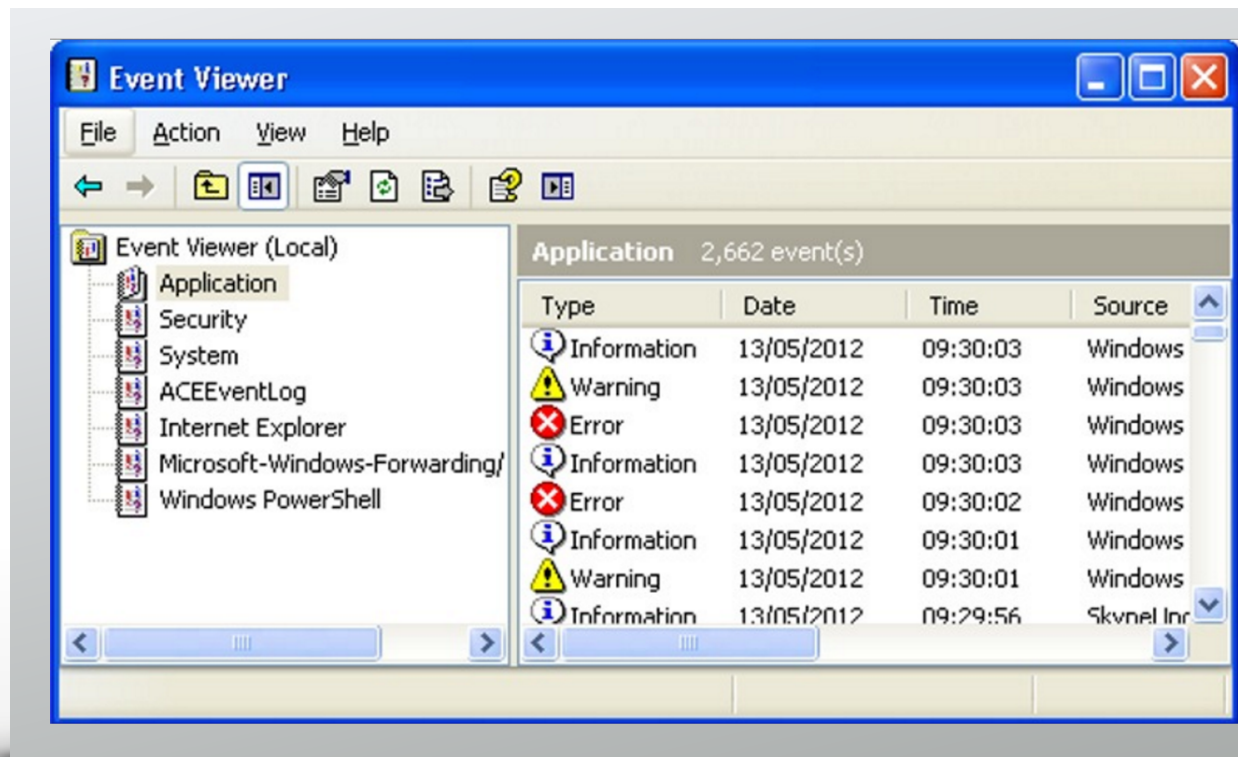
Packetbeat

Capture the Packet

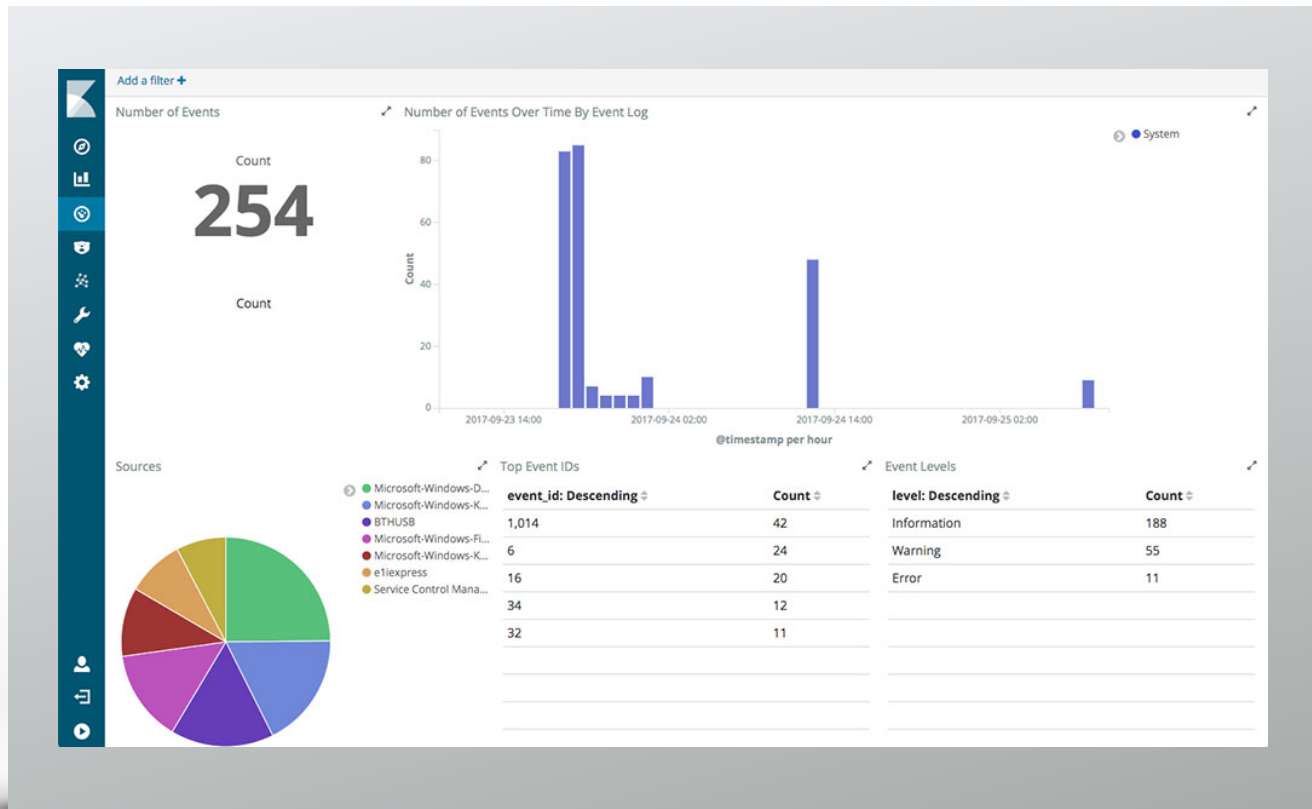


winlogbeat

Welcome
to **1998**



Now



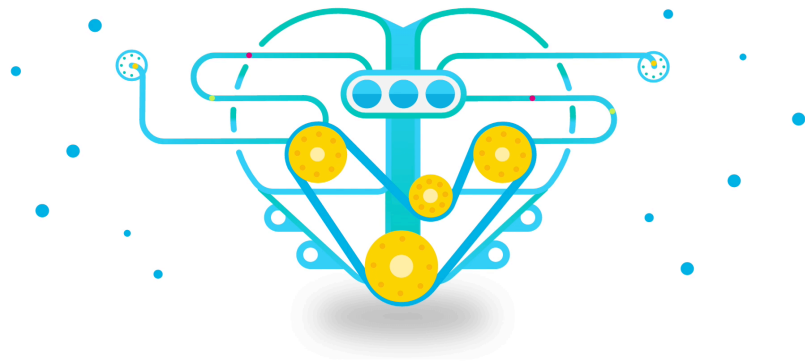


elasticsearch



Elasticsearch

Heart of the Elastic Stack



分散型、スケーラブル

高可用性

マルチテナント

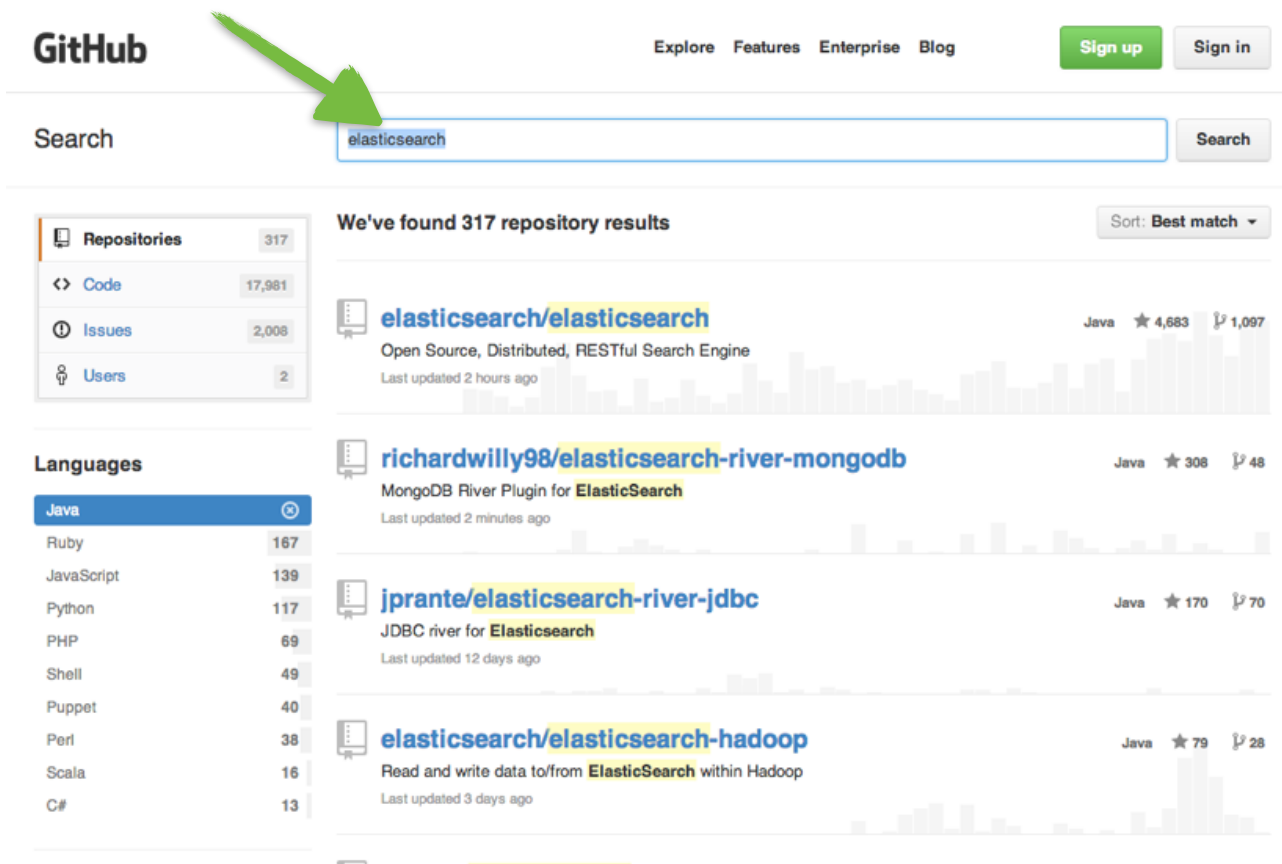
開発者フレンドリー

リアルタイム、全文検索

アグリゲーション

Elasticsearchとは？

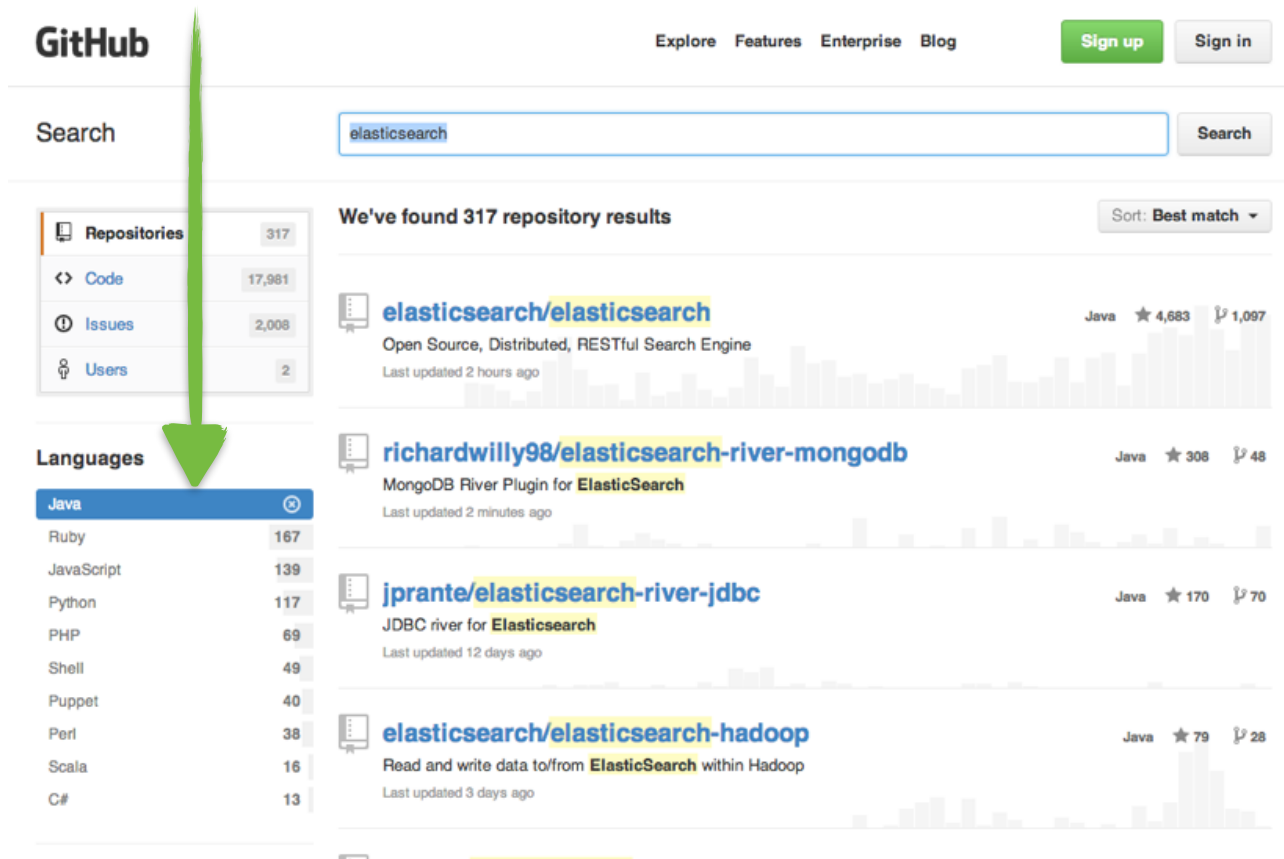
フリーワード検索



The screenshot shows the GitHub search interface. At the top left is the GitHub logo. To its right are links for 'Explore', 'Features', 'Enterprise', and 'Blog'. Further right are 'Sign up' and 'Sign in' buttons. Below the logo is a search bar containing the text 'elasticsearch', with a green arrow pointing to it from the left. To the right of the search bar is a 'Search' button. Below the search bar, the results are displayed under the heading 'We've found 317 repository results'. On the left side, there are filters for 'Repositories' (317), 'Code' (17,981), 'Issues' (2,008), and 'Users' (2). Below these is a 'Languages' section with a dropdown menu set to 'Java'. The search results list several repositories, each with a repository icon, the repository name, a description, the language, star count, and fork count. A bar chart is visible below each repository name.

Repository	Description	Language	Stars	Forks
elasticsearch/elasticsearch	Open Source, Distributed, RESTful Search Engine	Java	4,683	1,097
richardwilly98/elasticsearch-river-mongodb	MongoDB River Plugin for ElasticSearch	Java	308	48
jprante/elasticsearch-river-jdbc	JDBC river for Elasticsearch	Java	170	70
elasticsearch/elasticsearch-hadoop	Read and write data to/from ElasticSearch within Hadoop	Java	79	28

絞り込み



The screenshot shows the GitHub search interface. At the top, the GitHub logo is on the left, and navigation links for 'Explore', 'Features', 'Enterprise', and 'Blog' are in the center. On the right, there are 'Sign up' and 'Sign in' buttons. Below the navigation, the search bar contains the text 'elasticsearch' and a 'Search' button. A green arrow points from the search bar down to the 'Java' filter in the 'Languages' section on the left. The search results are displayed in a list format, showing repository names, descriptions, and star counts. The first result is 'elasticsearch/elasticsearch' with 4,683 stars. The second is 'richardwilly98/elasticsearch-river-mongodb' with 308 stars. The third is 'jprante/elasticsearch-river-jdbc' with 170 stars. The fourth is 'elasticsearch/elasticsearch-hadoop' with 79 stars. The 'Languages' section on the left lists various programming languages with their respective repository counts: Java (317), Code (17,981), Issues (2,008), Users (2), Ruby (167), JavaScript (139), Python (117), PHP (69), Shell (49), Puppet (40), Perl (38), Scala (16), and C# (13).

GitHub

Explore Features Enterprise Blog

Sign up Sign in

Search

elasticsearch Search

We've found 317 repository results Sort: Best match

Repositories 317

Code 17,981

Issues 2,008

Users 2

Languages

Java 317

Ruby 167

JavaScript 139

Python 117

PHP 69

Shell 49

Puppet 40

Perl 38

Scala 16

C# 13

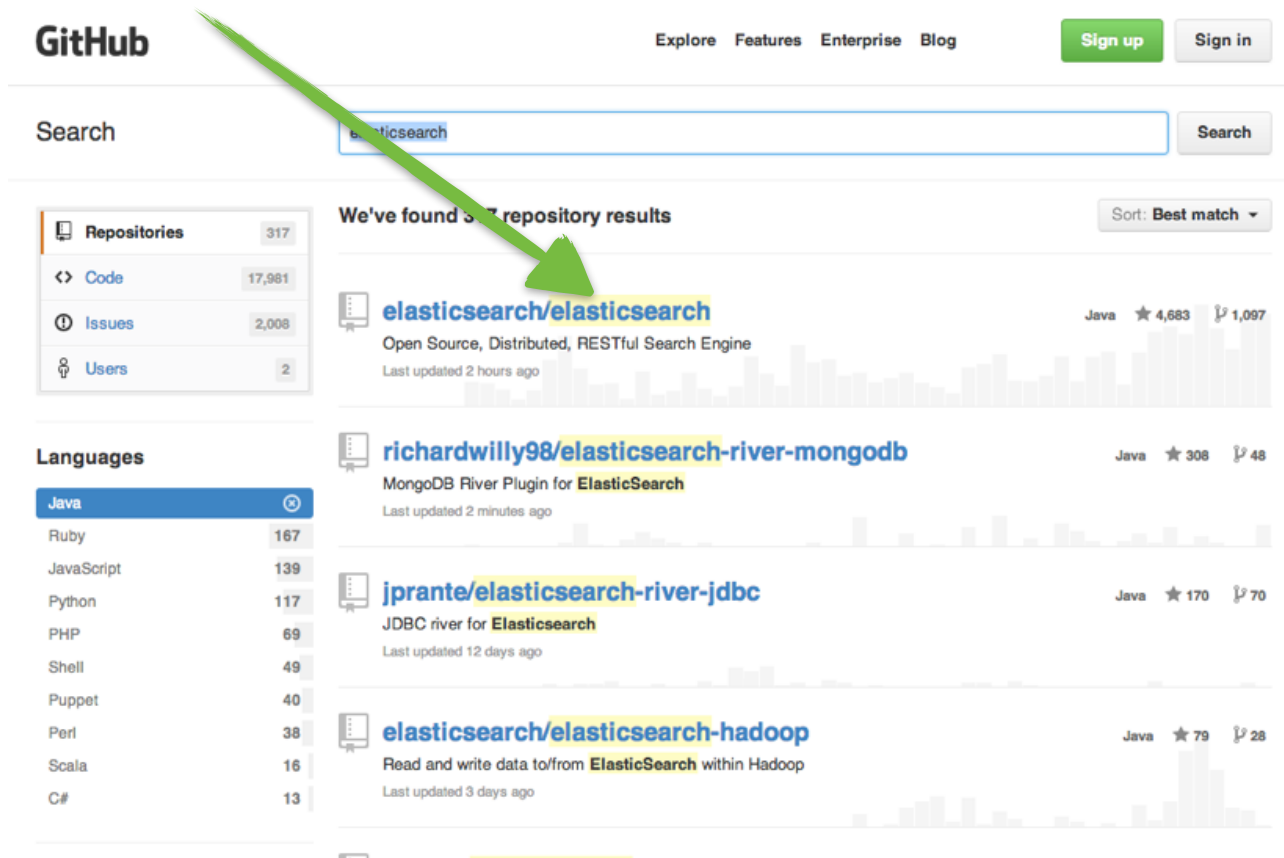
elasticsearch/elasticsearch Java ★ 4,683 1,097
Open Source, Distributed, RESTful Search Engine
Last updated 2 hours ago

richardwilly98/elasticsearch-river-mongodb Java ★ 308 48
MongoDB River Plugin for **ElasticSearch**
Last updated 2 minutes ago

jprante/elasticsearch-river-jdbc Java ★ 170 70
JDBC river for **Elasticsearch**
Last updated 12 days ago

elasticsearch/elasticsearch-hadoop Java ★ 79 28
Read and write data to/from **ElasticSearch** within Hadoop
Last updated 3 days ago

ハイライト



The screenshot shows the GitHub search interface. At the top, the GitHub logo is on the left, and navigation links for 'Explore', 'Features', 'Enterprise', and 'Blog' are in the center. On the right, there are 'Sign up' and 'Sign in' buttons. Below the navigation is a search bar containing the text 'elasticsearch' and a 'Search' button. To the left of the search results is a sidebar with filters: 'Repositories' (317), 'Code' (17,981), 'Issues' (2,008), and 'Users' (2). Below the filters is a 'Languages' section with a list of programming languages and their counts: Java (317), Ruby (167), JavaScript (139), Python (117), PHP (69), Shell (49), Puppet (40), Perl (38), Scala (16), and C# (13). The main search results area displays 'We've found 3 repository results' with a 'Sort: Best match' dropdown. The first result is 'elasticsearch/elasticsearch', which is highlighted in yellow. A green arrow points from the top left towards this result. The second result is 'richardwilly98/elasticsearch-river-mongodb' and the third is 'jprante/elasticsearch-river-jdbc'. Each result includes the repository name, a brief description, the last updated time, and a bar chart showing activity over time.

GitHub

Explore Features Enterprise Blog

Sign up Sign in

Search

elasticsearch Search

Sort: Best match

We've found 3 repository results

Repositories 317

Code 17,981

Issues 2,008

Users 2

Languages

Java 317

Ruby 167

JavaScript 139

Python 117

PHP 69

Shell 49

Puppet 40

Perl 38

Scala 16

C# 13

elasticsearch/elasticsearch Java ★ 4,683 1,097

Open Source, Distributed, RESTful Search Engine

Last updated 2 hours ago

richardwilly98/elasticsearch-river-mongodb Java ★ 308 48

MongoDB River Plugin for ElasticSearch

Last updated 2 minutes ago

jprante/elasticsearch-river-jdbc Java ★ 170 70

JDBC river for Elasticsearch

Last updated 12 days ago

elasticsearch/elasticsearch-hadoop Java ★ 79 28

Read and write data to/from ElasticSearch within Hadoop

Last updated 3 days ago

ソート

GitHub

Explore Features Enterprise Blog

Sign up

Sign in

Search

elasticsearch

Search

Repositories	317
Code	17,981
Issues	2,008
Users	2

Languages

Java	167
Ruby	139
JavaScript	117
Python	69
PHP	49
Shell	40
Puppet	38
Perl	16
Scala	13
C#	

We've found 317 repository results

Sort: Best match

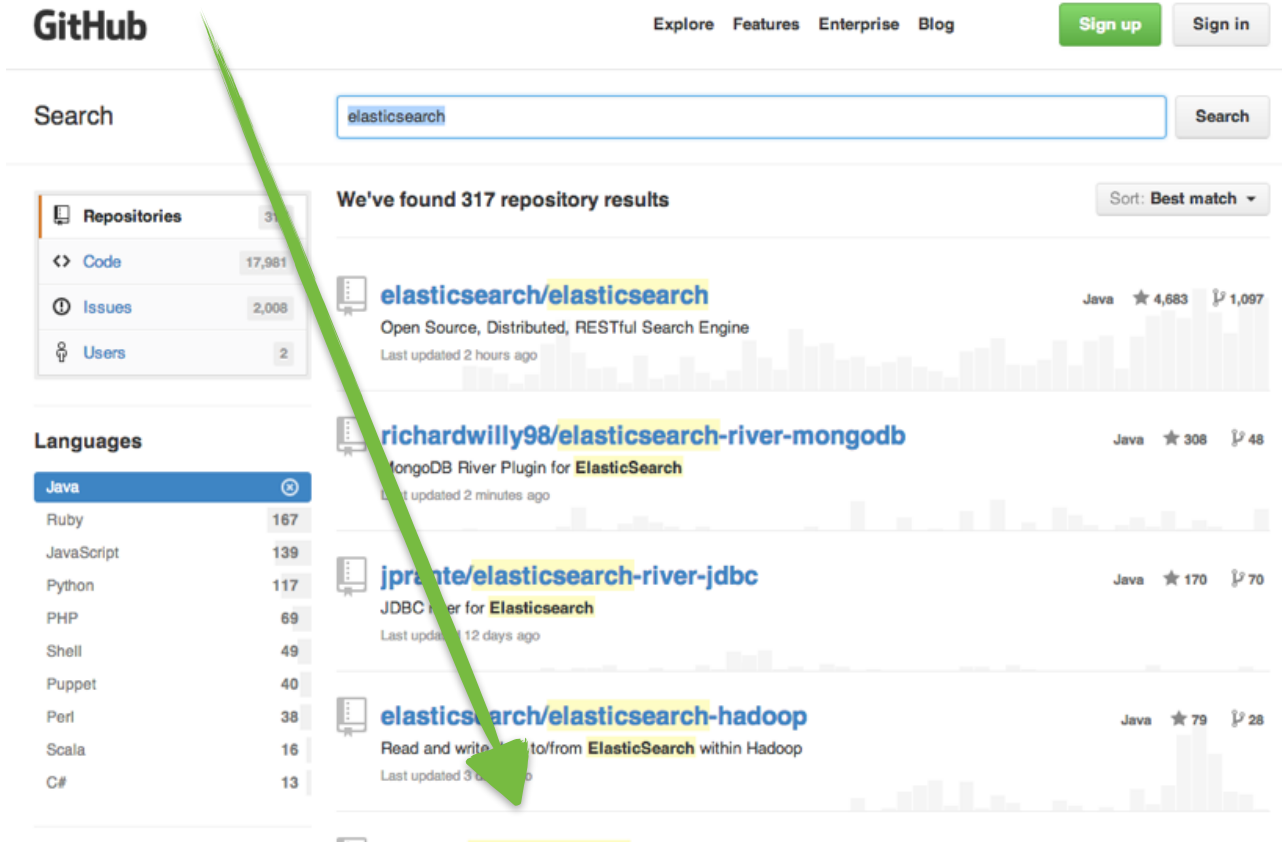
elasticsearch/elasticsearch Java ★ 4,683 1,097
Open Source, Distributed, RESTful Search Engine
Last updated 2 hours ago

richardwilly98/elasticsearch-river-mongodb Java ★ 308 48
MongoDB River Plugin for **ElasticSearch**
Last updated 2 minutes ago

jprante/elasticsearch-river-jdbc Java ★ 170 70
JDBC river for **Elasticsearch**
Last updated 12 days ago

elasticsearch/elasticsearch-hadoop Java ★ 79 28
Read and write data to/from **ElasticSearch** within Hadoop
Last updated 3 days ago

ページング



The screenshot shows the GitHub search interface. At the top, the GitHub logo is on the left, and navigation links for 'Explore', 'Features', 'Enterprise', and 'Blog' are in the center. On the right, there are 'Sign up' and 'Sign in' buttons. Below the navigation is a search bar containing the text 'elasticsearch' and a 'Search' button. To the left of the search results is a sidebar with filters: 'Repositories' (317), 'Code' (17,981), 'Issues' (2,008), and 'Users' (2). Below the filters is a 'Languages' section with a list of programming languages and their counts: Java (317), Ruby (167), JavaScript (139), Python (117), PHP (69), Shell (49), Puppet (40), Perl (38), Scala (16), and C# (13). The main search results area displays 'We've found 317 repository results' and a 'Sort: Best match' dropdown. The first result is 'elasticsearch/elasticsearch', an Open Source, Distributed, RESTful Search Engine, last updated 2 hours ago, with 4,683 stars and 1,097 forks. A green arrow points from the top left towards this first result.

GitHub

Explore Features Enterprise Blog

Sign up Sign in

Search

elasticsearch Search

We've found 317 repository results

Sort: Best match

Repositories 317

Code 17,981

Issues 2,008

Users 2

Languages

Java 317

Ruby 167

JavaScript 139

Python 117

PHP 69

Shell 49

Puppet 40

Perl 38

Scala 16

C# 13

elasticsearch/elasticsearch

Open Source, Distributed, RESTful Search Engine

Last updated 2 hours ago

Java 4,683 1,097

richardwilly98/elasticsearch-river-mongodb

MongoDB River Plugin for ElasticSearch

Last updated 2 minutes ago

Java 308 48

jprante/elasticsearch-river-jdbc

JDBC River for Elasticsearch

Last updated 12 days ago

Java 170 70

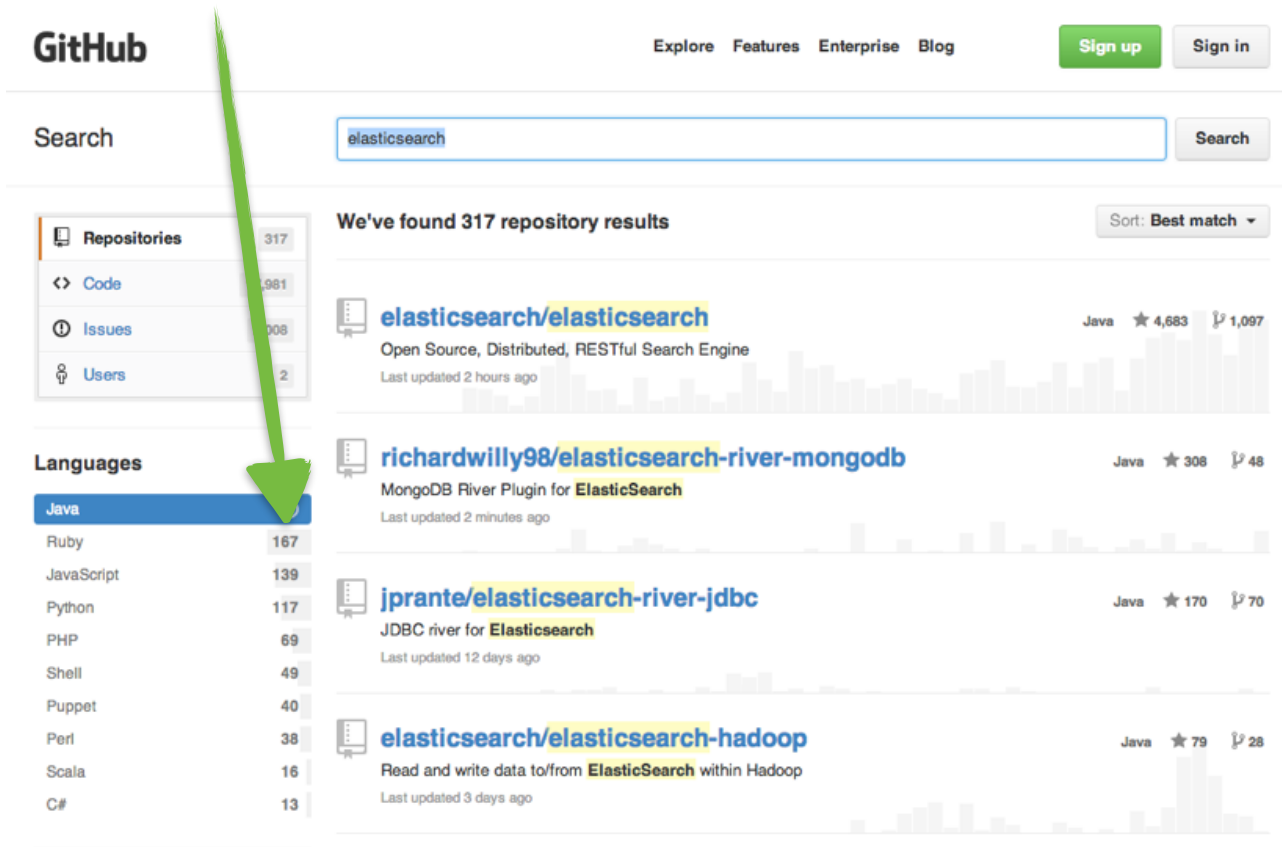
elasticsearch/elasticsearch-hadoop

Read and write to/from ElasticSearch within Hadoop

Last updated 3 days ago

Java 79 28

集計



The screenshot shows the GitHub search interface. At the top, the GitHub logo is on the left, and navigation links for 'Explore', 'Features', 'Enterprise', and 'Blog' are in the center. On the right, there are 'Sign up' and 'Sign in' buttons. Below the navigation is a search bar containing the text 'elasticsearch' and a 'Search' button. A green arrow points from the search bar down to the 'Java' language filter in the 'Languages' section on the left.

Search

Repositories 317
Code 981
Issues 008
Users 2

Languages
Java 167
Ruby 167
JavaScript 139
Python 117
PHP 69
Shell 49
Puppet 40
Perl 38
Scala 16
C# 13

We've found 317 repository results Sort: Best match

- elasticsearch/elasticsearch** Java ★ 4,683 1,097
Open Source, Distributed, RESTful Search Engine
Last updated 2 hours ago
- richardwilly98/elasticsearch-river-mongodb** Java ★ 308 48
MongoDB River Plugin for **ElasticSearch**
Last updated 2 minutes ago
- jprante/elasticsearch-river-jdbc** Java ★ 170 70
JDBC river for **Elasticsearch**
Last updated 12 days ago
- elasticsearch/elasticsearch-hadoop** Java ★ 79 28
Read and write data to/from **ElasticSearch** within Hadoop
Last updated 3 days ago

サジェスト

The screenshot shows the GitHub repository page for `elasticsearch/elasticsearch`. A search bar at the top right contains the text `repository debian`. A green arrow points to this search bar. A dropdown menu is open below the search bar, displaying several search suggestions:

- elasticsearch/elasticsearch#1726 debian package violates naming convention
- elasticsearch/elasticsearch#3571 debian package init-script: start-stop-daemon ne
- elasticsearch/elasticsearch#1681 Debian pkg
- elasticsearch/elasticsearch#3286 There is no official debian/ubuntu repository
- elasticsearch/elasticsearch#3500 Elasticsearch should include debian's standard j
- elasticsearch/elasticsearch#1526 Moving debian package to maven

Below the suggestions are two search options:

- Search elasticsearch/elasticsearch for 'debian'
- Search GitHub for 'debian'

The background of the page shows the repository's issue list. The left sidebar includes the GitHub logo, repository name, and navigation options like "Browse Issues" and "Everyone's Issues". The right sidebar shows "Sign up" and "Sign in" buttons, along with "Star" (4,683) and "Fork" (1,097) buttons. The main content area displays a list of issues, including:

- NoShardAvailableActionException in ES 0.90.3 on startup** (Issue #3700)
- Feature Request: Don't reindex the document when updating non-indexed fields** (Issue #3696)

Labels on the left include Lucene 4.5 Upgrade, breaking, bug, enhancement, feature, and non-issue.

Elasticsearch in 10 seconds

- スキーマフリー、分散ドキュメントストア、REST & JSON
- オープンソース: Apache License 2.0
- 設定なしで簡単に試すことが可能
- Javaで実装。拡張も容易

簡単なCRUD

データ登録

```
curl -XPUT localhost:9200/books/book/1 -d '{
  "title" : "Elasticsearch – The definitive guide",
  "authors" : "Clinton Gormley",
  "started" : "2013-02-04",
  "pages" : 230
}'
```

データ更新

```
curl -XPUT localhost:9200/books/book/1 -d '{
  "title" : "Elasticsearch - The definitive guide",
  "authors" : [ "Clinton Gormley", "Zachary Tong" ],
  "started" : "2013-02-04",
  "pages" : 230
}'
```

データ削除

```
curl -X DELETE localhost:9200/books/book/1
```

データの取得

```
curl -X GET localhost:9200/books/book/1
```

```
curl -X GET localhost:9200/books/book/1/_source
```

検索 - Query DSL

```
curl -XGET 'localhost:9200/books/doc/_search' -d '{
  "query": {
    "bool": {
      "must": [
        { "match": { "title": "Search" }},
        { "match": { "content": "Elasticsearch" }}
      ],
      "filter": [
        { "term": { "status": "published" }},
        { "range": { "publish_date": { "gte": "2015-01-01" }}}
      ]
    }
  }
}'
```

分散構成、 スケール

Basic terms

- インデックス
 - データの論理的な集合。
RDBのデータベースのようなものLogical
- レプリケーション
 - 読み込みのスケラビリティ向上
 - SPOFの解消
- シャーディング
 - 複数マシンへデータを分割
書き込みのスケラビリティ向上
データフロー制御

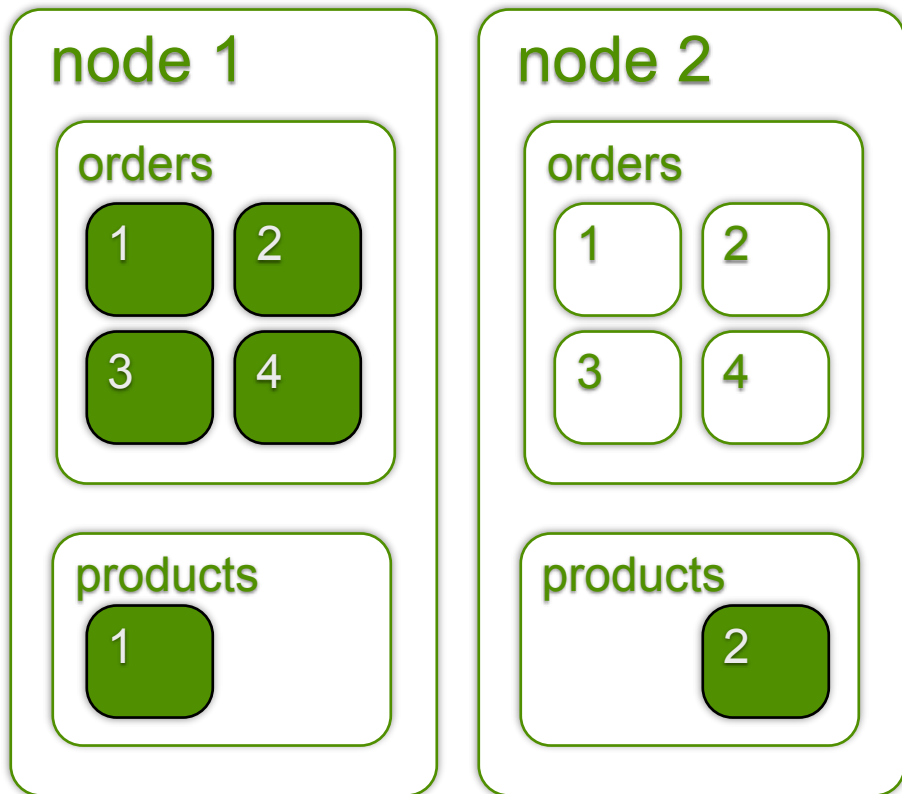
シャードとレプリカ



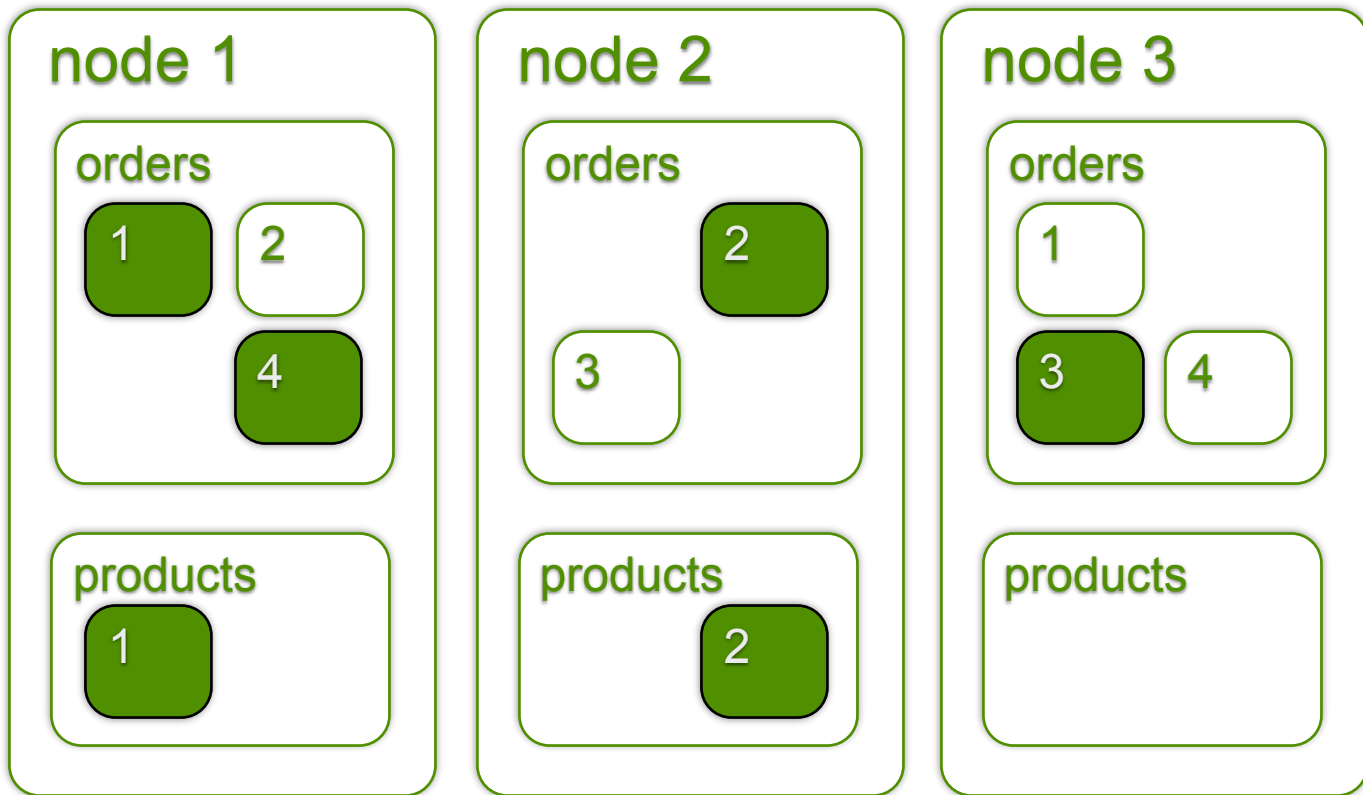
```
curl -X PUT localhost:9200/orders -d '{  
  "settings.index.number_of_shards" : 4  
  "settings.index.number_of_replicas" : 1  
}'
```

```
curl -X PUT localhost:9200/products -d '{  
  "settings.index.number_of_shards" : 2  
  "settings.index.number_of_replicas" : 0  
}'
```

シャードとレプリカ



自動的な分散



全文検索とは？

全文検索とは？

- 全文検索 (Full text search) とは、コンピュータにおいて、複数の文書 (ファイル) から特定の文字列を検索すること。「ファイル名検索」や「単一ファイル内の文字列検索」と異なり、「複数文書にまたがって、文書に含まれる全文を対象とした検索」という意味で使用される。
(Wikipediaより)

用語

- インデックス
 - 検索エンジンが検索に使用するデータの保存先
- ドキュメント（文書）
 - 検索エンジンに保存されたデータ
- フィールド
 - ドキュメントに含まれる属性
- クエリ
 - 検索条件、検索式

用語

- スキーマ
 - ドキュメントの構造を定義するもの
- ターム (Term)、トークン (Token)
 - インデックスのキーになる単語 (文字列)
 - 文章を一定の法則で区切った単語
 - 単語だけでなく、単語の位置なども含む

ドキュメントの登録

1 カツオはサザエの弟

2 サザエはワカメの姉

ドキュメントの登録

ドキュメントの登録

1

カツオはサザエの弟

2

サザエはワカメの姉

ドキュメントの登録

1

カツオ は サザエ の 弟

2

サザエ は ワカメ の 姉

単語に分割

ドキュメントの登録

1 カツオはサザエの弟

2 サザエはワカメの姉

ドキュメントの登録

1 カツオ は サザエ の 弟

2 サザエ は ワカメ の 姉

単語に分割

カツオ	1	の	1 2	弟	1
サザエ	1 2	は	1 2	姉	2
ワカメ	2				

単語からidの配列が引けるように

検索

カツオ サザエ

検索条件入力

カツオ	1	の	1 2	弟	1
サザエ	1 2	は	1 2	姉	2
ワカメ	2				

検索

カツオ サザエ

検索条件入力

カツオ AND サザエ

検索条件のパーズ
検索クエリ化

カツオ	1	の	1 2	弟	1
サザエ	1 2	は	1 2	姉	2
ワカメ	2				

検索

カツオ サザエ

検索条件入力

カツオ

AND

サザエ

検索条件のパーズ
検索クエリ化

カツオ

1

の

1

2

弟

1

サザエ

1

2

は

1

2

姉

2

ワカメ

2

検索

カツオ サザエ

検索条件入力

カツオ

AND

サザエ

検索条件のパーズ
検索クエリ化

カツオ

1

の

1

2

弟

1

サザエ

1

2

は

1

2

姉

2

ワカメ

2

検索

カツオ サザエ

検索条件入力

カツオ

AND

サザエ

検索条件のパーズ
検索クエリ化

カツオ

1

の

1

2

弟

1

サザエ

1

2

は

1

2

姉

2

ワカメ

2

検索

カツオ サザエ

検索条件入力

カツオ

AND

サザエ

検索条件のパーズ
検索クエリ化

カツオ

1

の

1

2

弟

1

サザエ

1

2

は

1

2

姉

2

ワカメ

2

検索

カツオ サザエ

検索条件入力

カツオ

AND

サザエ

検索条件のパーズ
検索クエリ化

カツオ

1

の

1

2

弟

1

サザエ

1

2

は

1

2

姉

2

ワカメ

2

単語の区切り方

- 英語の場合

I am speaking Introduction Elasticsearch.

- 日本語の場合

私は入門Elasticsearchについて話している。

単語の区切り方

- 英語の場合

I am speaking Introduction Elasticsearch.



スペースが切れ目とわかる

- 日本語の場合

私は入門Elasticsearchについて話している。



どこで区切れればよい？

N-Gramと形態素解析

- 転置インデックスのキーの作り方
 - 日本語は単語の切れ目がわからないので、転置インデックスのキーは主に次の2つの手法で作成
- N-Gram
 - N文字ずつ文章を区切る
- 形態素解析
 - 辞書などを用いて意味のある単語で区切る

形態素解析

カツオはサザエの弟

カツオ は サザエ の 弟

- メリット：
 - 意味のある単語の切れ目
品詞情報を元に追加処理が可能（語幹変換など）
- デメリット：
 - 新語（未知語）に弱い→辞書ベースの場合、辞書にない単語は検出不能。

N-Gram

カツオはサザエの弟

カツ

ツオ

オは

はサ

サザ

ザエ

エの

の弟

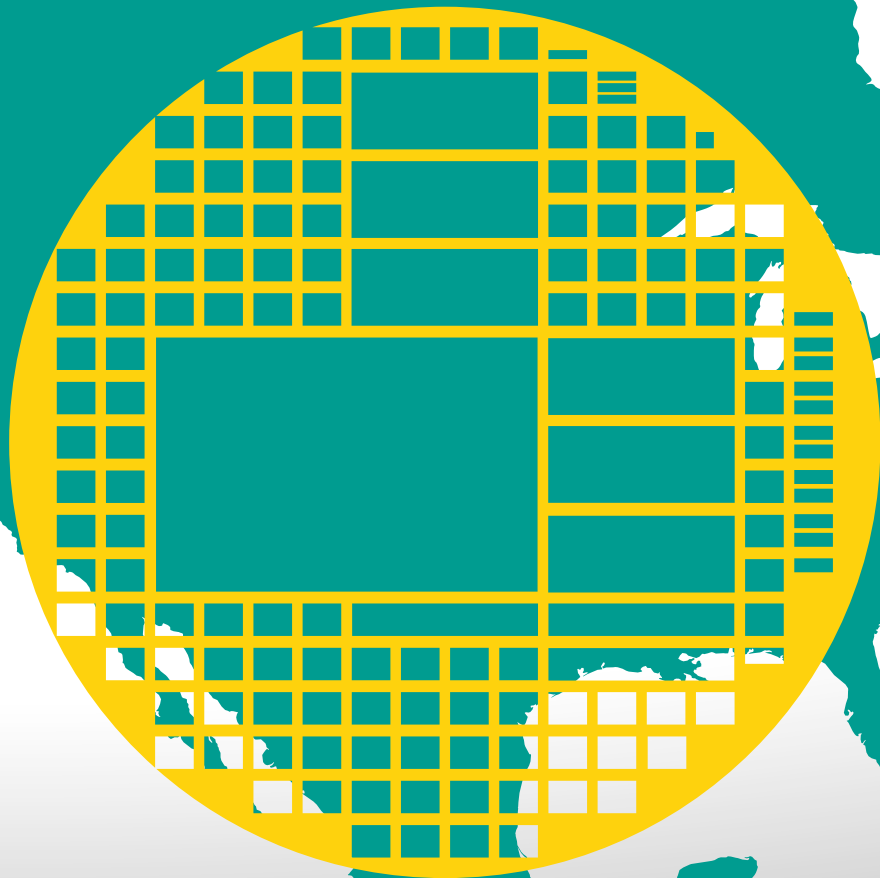
- メリット：
 - 未知語に対応可能
- デメリット：
 - インデックス肥大化
 - 品詞情報に基づく処理が不可能

その他の機能

GEO

さまざまな形式のデータで
Geo検索可能

緯度経度、GeoHash、
GeoShape...



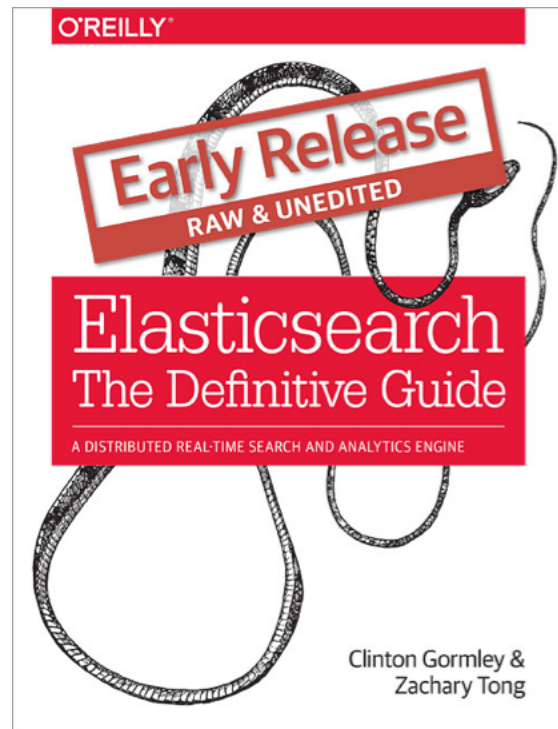
Ecosystem

- Plugins
 - プラグインによる機能の追加
- クライアントライブラリ
 - Java, Ruby, python, php, perl, javascript, .NET
 - Scala, clojure, go

詳しく知りたい方は

Elasticsearch – The Definitive guide

<http://www.elastic.co/guide/en/elasticsearch/guide/current/index.html>





kibana



Kibana

Window into the Elastic Stack



可視化と分析

地理空間

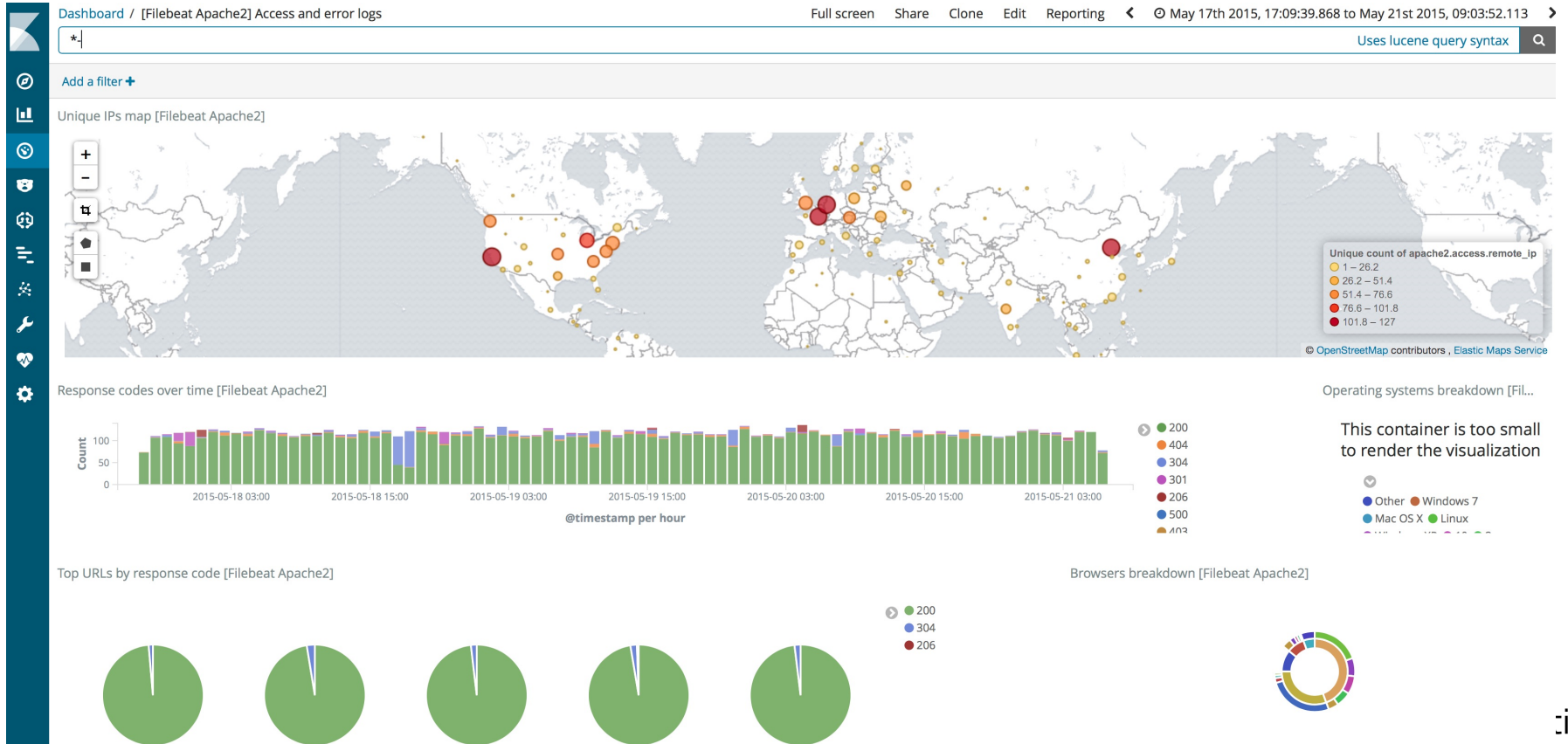
カスタマイズと
レポートの共有

グラフ探索

Elastic Stackへの
セキュアなアクセスと管理

カスタムAppsの作成

Kibana 6

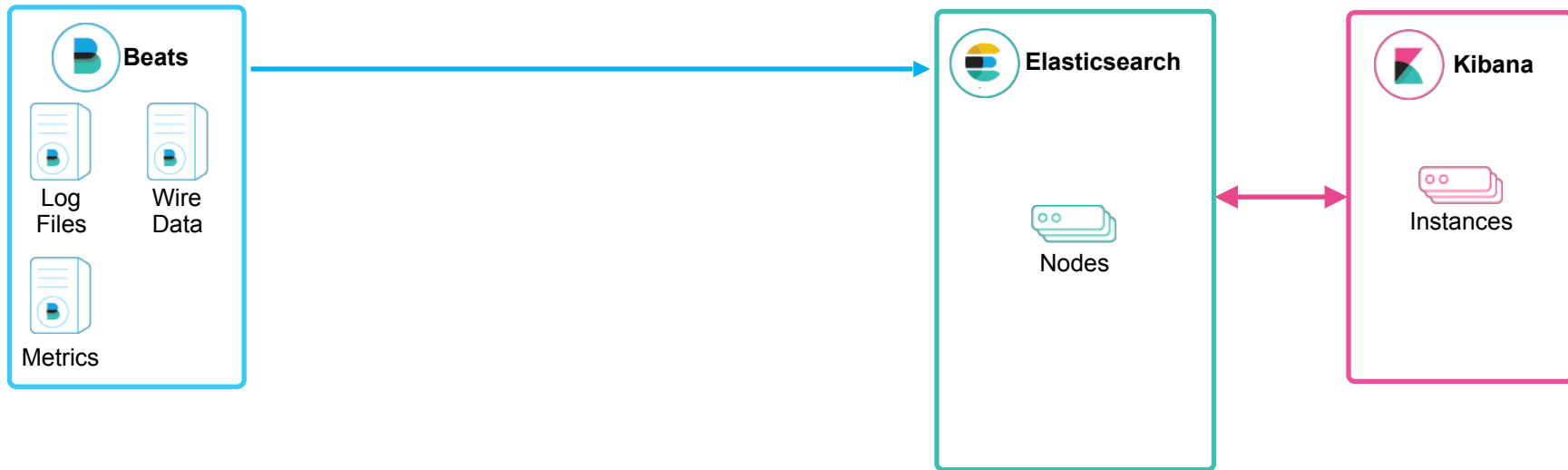


デモ

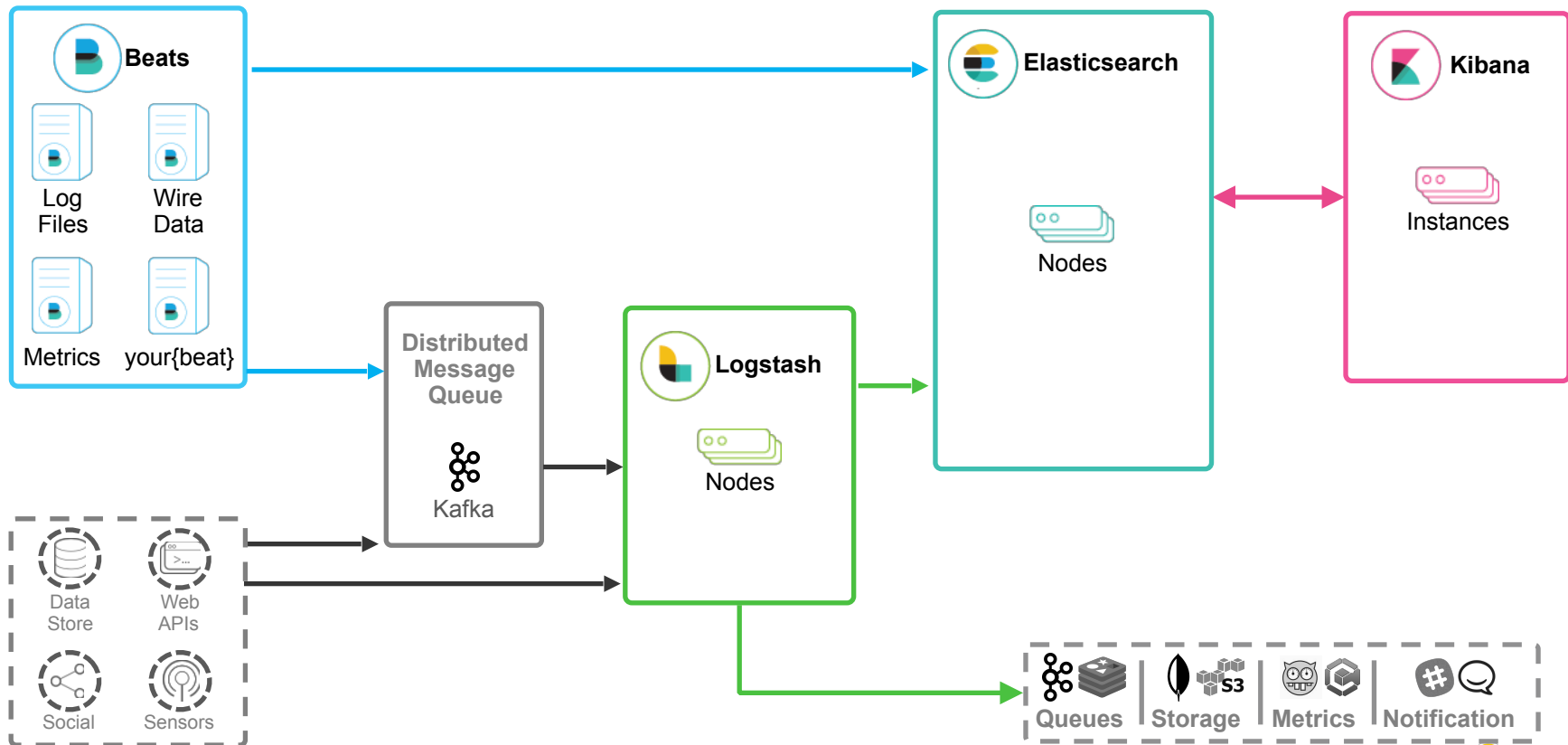
データ投入から可視化まで

本格的に解析を行うには？

Elastic Stackの構成



Elastic Stackの構成



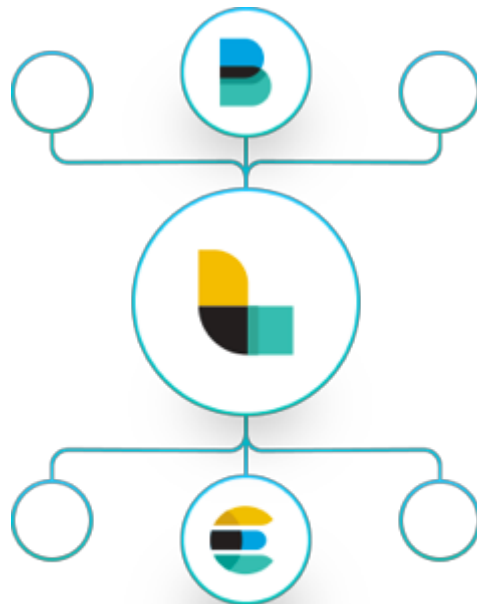


logstash



Logstash

データ加工パイプライン



全ての形式、サイズとデータソースの投入

安全で暗号化されたデータ入力

パースと動的なデータ変換

独自のパイプライン処理の作成

あらゆる出力にデータ転送

200以上のプラグイン

Logstash in 10 seconds

- ログ・データの収集・管理
- 収集、パース・加工、送出
- オープンソース：Apache License 2.0
- Ruby app (JRuby)



Logstash architecture

Input

collect and split

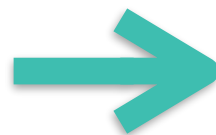
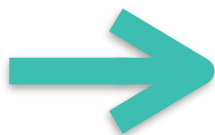
Filter

alter and enrich

Output

store and visualize

?



?

logstash

設定

```
input {  
  ...  
}  
  
filter {  
  ...  
}  
  
output {  
  ...  
}
```

設定：input

```
input {
  file {
    path => "/Users/johtani/sample/*_log"
    start_position => "beginning"
  }
}
```


1行1データ



```
189.120.xx.xx - - [02/Dec/2014:12:18:29 +0900] "GET /manager/html HTTP/1.1" 404 274 "-" "Mozilla/5.0 (Windows NT 5.1; rv:5.0) Gecko/20100101 Firefox/5.0"
```

設定：filter

```
filter {  
  grok {  
    match => { "message" => "%{COMBINEDAPACHELOG}" }  
    break_on_match => false  
  }  
  date {  
    match => ["timestamp", "dd/MMM/YYYY:HH:mm:ss Z"]  
    locale => en  
  }  
  geoip { source => ["clientip"] }  
  useragent {  
    source => "agent"  
    target => "useragent"  
  }  
}
```

パース

```
189.120.xx.xx - - [02/Dec/2014:12:18:29 +0900] "GET /manager/html HTTP/1.1"
404 274 "-" "Mozilla/5.0 (Windows NT 5.1; rv:5.0) Gecko/20100101 Firefox/5.0"
```



```
{...
  "@timestamp": "2015-04-10T09:07:49.325Z",
  "clientip": "189.120.xx.xx",
  "ident": "-",
  "auth": "-",
  "timestamp": "02/Dec/2014:12:18:29 +0900",
  "verb": "GET",
  "request": "/manager/html",
  "...
  "agent": "\"Mozilla/5.0 (Windows NT 5.1; rv:5.0) Gecko/elastic
```

設定：filter

```
filter {
  grok {
    match => { "message" => "%{COMBINEDAPACHELOG}" }
    break_on_match => false
  }
  date {
    match => ["timestamp", "dd/MMM/YYYY:HH:mm:ss Z"]
    locale => en
  }
  geoip { source => ["clientip"] }
  useragent {
    source => "agent"
    target => "useragent"
  }
}
```

日付のパーズ

```
{...  
  "@timestamp": "2015-04-10T09:07:49.325Z",  
  ...  
  "timestamp": "02/Dec/2014:12:18:29 +0900",  
  ...  
}
```



```
{...  
  "@timestamp": "2014-12-02T03:18:29.000Z",  
  ...  
  "timestamp": "02/Dec/2014:12:18:29 +0900",  
  ...  
}
```

設定：filter

```
filter {
  grok {
    match => { "message" => "%{COMBINEDAPACHELOG}" }
    break_on_match => false
  }
  date {
    match => ["timestamp", "dd/MMM/YYYY:HH:mm:ss Z"]
    locale => en
  }
  geolocation { source => ["clientip"] }
  useragent {
    source => "agent"
    target => "useragent"
  }
}
```

IPから緯度経度など付与

```
"clientip": "189.120.xx.xx",
```



```
"clientip": "189.120.xx.xx",  
"geoip": {  
  "ip": "189.120.xxx.xxx",  
  ...  
  "country_name": "Brazil",  
  "continent_code": "SA",  
  "region_name": "27",  
  "city_name": "São Paulo",  
  "latitude":
```

設定：filter

```
filter {
  grok {
    match => { "message" => "%{COMBINEDAPACHELOG}" }
    break_on_match => false
  }
  date {
    match => ["timestamp", "dd/MMM/YYYY:HH:mm:ss Z"]
    locale => en
  }
  geoip { source => ["clientip"] }
  useragent {
    source => "agent"
    target => "useragent"
  }
}
```


ユーザエージェントのパーズ

```
"agent": "\"Mozilla/5.0 (Windows NT 5.1; rv:5.0) Gecko/20100101 Firefox/5.0\""
```



```
"agent": "\"Mozilla/5.0 (Windows NT 5.1; rv:5.0) Gecko/20100101 Firefox/5.0\""  
  "useragent": {  
    "name": "Firefox",  
    "os": "Windows XP",  
    "os_name": "Windows XP",  
    "device": "Other",  
    "major": "5",  
    "minor": "0"
```

```
  }
```

設定 : output

```
output {
  elasticsearch {
    hosts => ["localhost"]
    index => "demo_access_log-%{+YYYY.MM.dd}"
  }
}
```

さらに活用するには？

elasticsearch-hadoop

ES-Hadoop : Hadoop、Spark、その他の製品との統合

- Hadoopデータのリアルタイムサーチ
- Hadoopに関するスタンドアロンの自己充足型のライブラリ
- ESデータへの双方向アクセス
- MapReduce、Hive、Pig、Cascading、Spark、Stormをサポート
- HDFSを使用してESデータのバックアップやアーカイブを実施



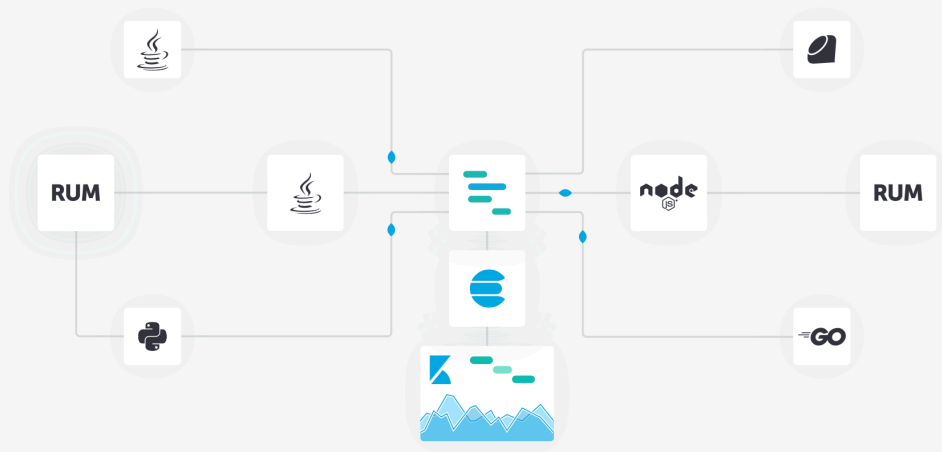


オープンソースのアプリケーション パフォーマンス監視 (APM)

ログやシステムのメトリックをElasticsearchに取り込みましたか？
ElasticのAPMで、アプリケーションのメトリックも取り込むことができます。

初期設定に、4行コードを加えるだけ。

問題箇所をすばやく確認し、自信をもってコードをプッシュできます。



ElasticのAPMで、パフォーマンスメトリックの可視化が簡単に。 | [今すぐトライ](#)

NEW Elastic APM UIに新メニューが登場。検索バー、機械学習統合、RubyとJavaScriptのRUM向けエージェント、JavaとGoのベータ版が加わりました。 [さらに詳しく](#)



ELASTIC STACK

Elastic Stackのオプション

エンタープライズグレードのセキュリティと、開発者フレンドリーなAPIを備えたオプション（旧X-Pack）。機械学習からグラフ分析まで、多彩な機能を手軽に、楽しく使えます。



セキュリティ

Elasticsearchデータを堅牢に、きめ細やかな設定で保護

[さらに詳しく](#)

アラート

データの変化を通知

[さらに詳しく](#)

監視

Elastic Stackを監視し、高水準な稼働状況を保つ

[さらに詳しく](#)



ELASTIC CLOUD

Elasticsearchのパワーを利用したSaaS製品群

Elastic Cloudは、展開、運用、スケールが容易にできるElasticの製品とソリューションをCloudで利用可能にした、成長し続けるSaaS製品群です。容易に利用できるElasticsearchのマネージドサービスから、パワフルですぐに利用可能なソリューションまで、Elastic Cloudは、Elasticを継ぎ目なく業務に適用するための足がかりです。



Elasticsearch Service

AWSやGCPで、Kibanaや他では得られない機能と共に容易に展開します。

製品概要

今すぐトライ



Elastic App Search Service

アプリケーションにスケーラブルな検索機能を実装するために、ものの数分で展開します。

製品概要

今すぐトライ



Elastic Site Search Service

パワフルな検索体験をあなたのウェブサイトで提供できます。特別な学習は必要ではありません。

製品概要

今すぐトライ

その他の使い方

データの登録方法

- Kibanaのサンプルデータ (6.4から)
- LogstashでJDBC input
- LogstashでCSV
- Filebeatでアクセスログ
- Metricbeatでメトリック
- PacketbeatでMySQL/PostgreSQLのパケット解析

Kibanaのサンプルデータ (>= 6.4.0)

The screenshot shows the Kibana interface with a sidebar on the left containing navigation items: Discover, Visualize, Dashboard, Timelion, Machine Learning, Graph, Dev Tools, Monitoring, and Management. The main content area is titled 'Add Data to Kibana' and includes three primary options: Logging, Metrics, and Security Analytics. Each option has a descriptive paragraph and a corresponding 'Add' button. Below these options, there are two links: 'Sample Data' (highlighted with a red box) and 'Your Data'. The 'Sample Data' link is described as 'Load a data set and a Kibana dashboard'. Below this section, there are two more main sections: 'Visualize and Explore Data' and 'Manage and Administer the Elastic Stack', each containing several sub-options with icons and brief descriptions.

Add Data to Kibana

Use these solutions to quickly turn your data into pre-built dashboards and monitoring systems.

- Logging**
Ingest logs from popular data sources and easily visualize in preconfigured dashboards.
[Add log data](#)
- Metrics**
Collect metrics from the operating system and services running on your servers.
[Add metric data](#)
- Security Analytics**
Centralize security events for interactive investigation in ready-to-go visualizations.
[Add security events](#)

Sample Data Load a data set and a Kibana dashboard

Your Data Connect to your Elasticsearch index

Visualize and Explore Data

- Dashboard**
Display and share a collection of visualizations and saved searches.
- Discover**
Interactively explore your data by querying and filtering raw documents.
- Graph**
Surface and analyze relevant relationships in your Elasticsearch data.
- Machine Learning**
Automatically model the normal behavior of your time series data to detect anomalies.
- Timelion**
Use an expression language
- Visualize**
Create visualizations and

Manage and Administer the Elastic Stack

- Console**
Skip cURL and use this JSON interface to work with your data directly.
- Index Patterns**
Manage the index patterns that help retrieve your data from Elasticsearch.
- Logstash Pipelines**
Create, delete, update, and clone data ingestion pipelines.
- Monitoring**
Track the real-time health and performance of your Elastic Stack.
- Saved Objects**
Import, export, and manage your saved searches.
- Security Settings**
Protect your data and easily manage who has access to

ワンクリックでデータ登録

Home

Add Data to Kibana

All Logging Metrics Security Analytics **Sample Data**

Sample flight data
Sample data, visualizations, and dashboards for monitoring flight routes.

[Add](#)

Management / Elasticsearch / Index Management

Index management

Update your Elasticsearch indices individually

flight

<input type="checkbox"/>	Na...	Health
<input type="checkbox"/>	kibana_sample_data_flights	● green

Rows per page: 10

kibana_sample_data_flights

[Summary](#) [Settings](#) [Mapping](#) [Stats](#) [Edit settings](#)

Health: ● green

Status: open

Primaries: 1

Replicas: 0

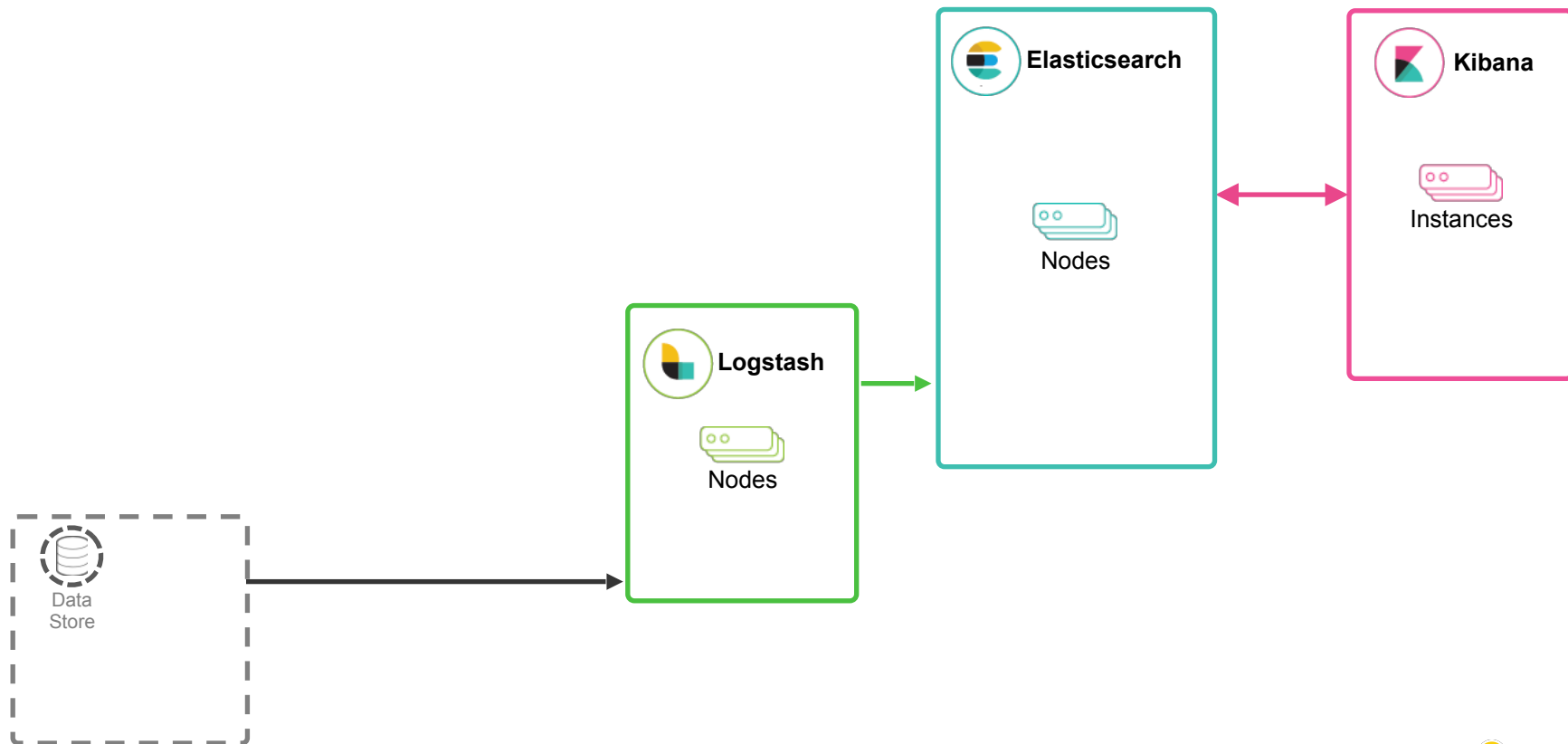
Docs Count: 13059

Docs Deleted: 0

Storage Size: 6.4mb

Primary Storage Size: 6.4mb

Logstash JDBC Input



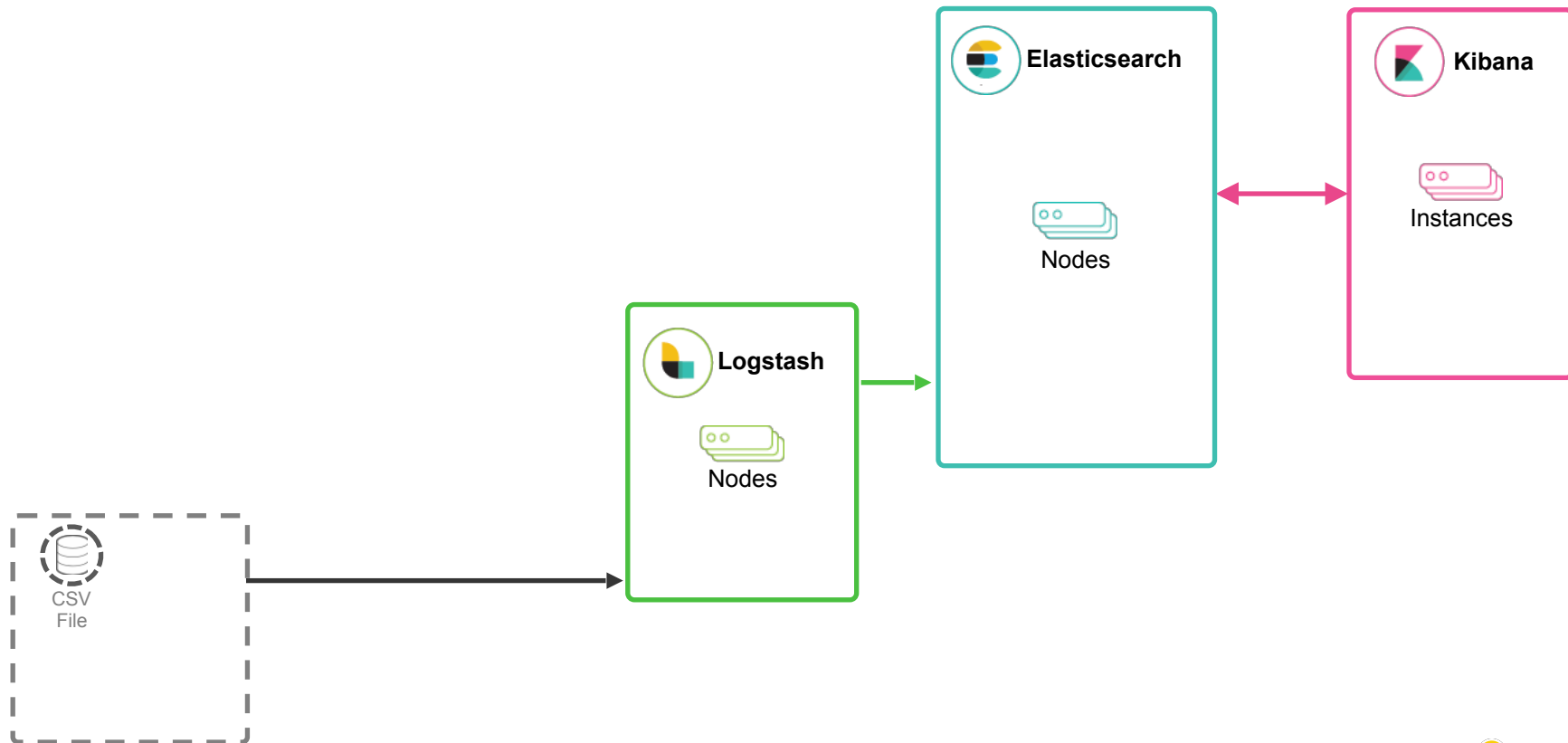
JDBC Input

```
input {
  jdbc {
    jdbc_connection_string => "jdbc:postgresql://db_server:5432/"
    jdbc_driver_class => "org.postgresql.Driver"
    jdbc_driver_library => "/home/elastic/postgresql-9.1-901-1.jdbc4.jar"
    jdbc_user => "postgres"
    jdbc_password => "password"
    statement => "SELECT * from blogs"
  }
}

filter {
  mutate {
    remove_field => ["@version", "host", "message", "@timestamp", "id", "tags"]
  }
}

output {
  stdout { codec => "dots" }
  elasticsearch {
    index => "blogs"
    document_type => "_doc"
  }
}
```

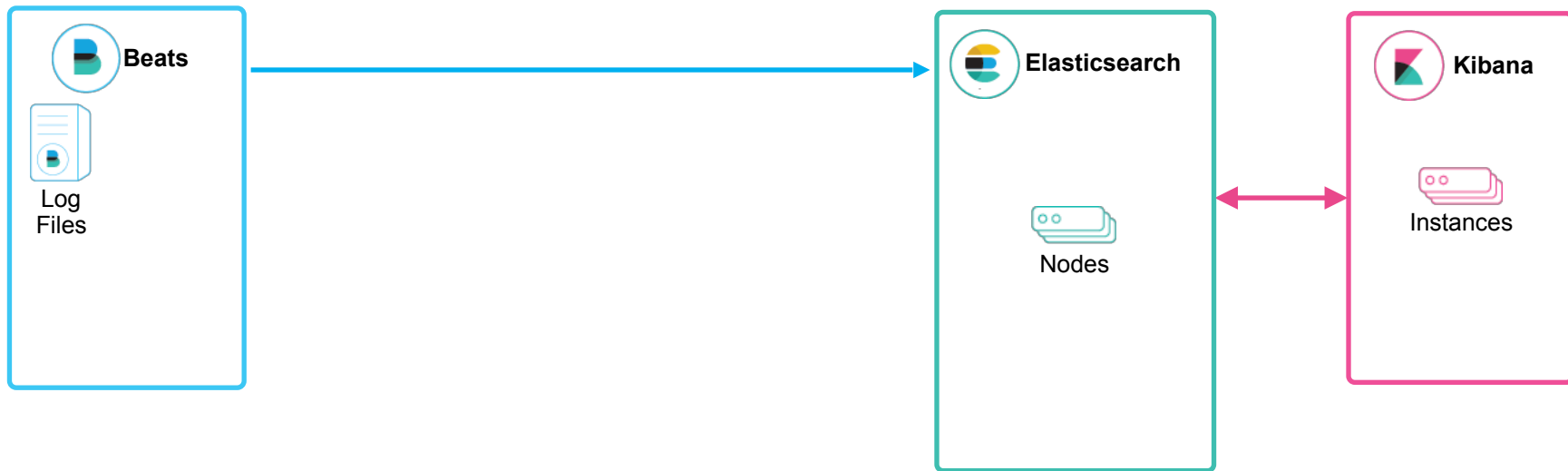
Logstash CSV



CSV filter

```
input {
  stdin { type => "${EVENT:earthquake}" }
}
filter {
  csv {
    columns => ["timestamp","latitude","longitude","depth","mag","magType","nst","gap","dmin","rms","source","event_id"]
    convert => {"latitude" => "float"}
    convert => {"longitude" => "float"}
    convert => {"depth" => "float"}
    convert => {"mag" => "float"}
    convert => {"dmin" => "float"}
    convert => {"rms" => "float"}
    convert => {"gap" => "float"}
  }
  mutate {
    add_field => ["location", "%{latitude}, %{longitude}"]
    remove_field => ["latitude", "longitude"]
  }
  date {
    match => ["timestamp", "yyyy/MM/dd HH:mm:ss.SS", "ISO8601"]
    remove_field => ["timestamp"]
    timezone => "GMT"
  }
}
output {
  # stdout { codec => rubydebug { metadata => true } }
  stdout { codec => dots }
  elasticsearch {
    # hosts => [""]
```

Filebeatでアクセスログ



Filebeatでアクセスログ

- 2つのElasticsearchプラグインをインストールしてElasticsearchを起動

```
bin/elasticsearch-plugin install ingest-geoip  
bin/elasticsearch-plugin install ingest-user-agent
```

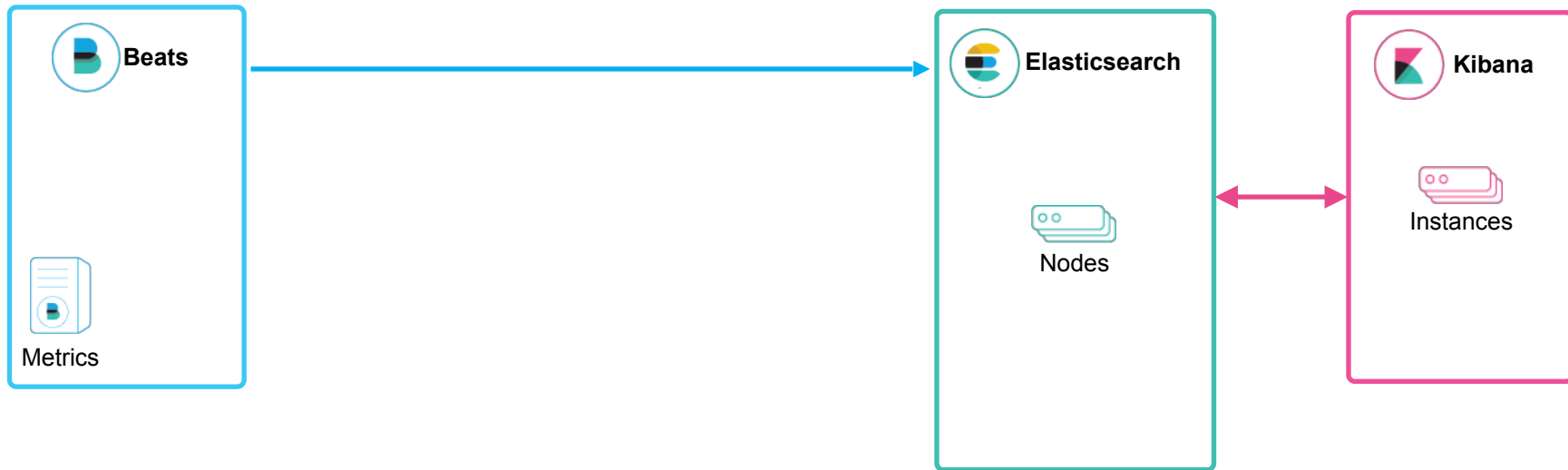
- Filebeatのapache2モジュールを有効化

```
./filebeat modules enable apache2
```

- modules.d/apache2.ymlにアクセスログのパスを設定
- setupコマンドを実行してからFilebeatを起動

```
./filebeat setup  
./filebeat -e
```

Metricbeatでメトリック



Metricbeatでメトリック

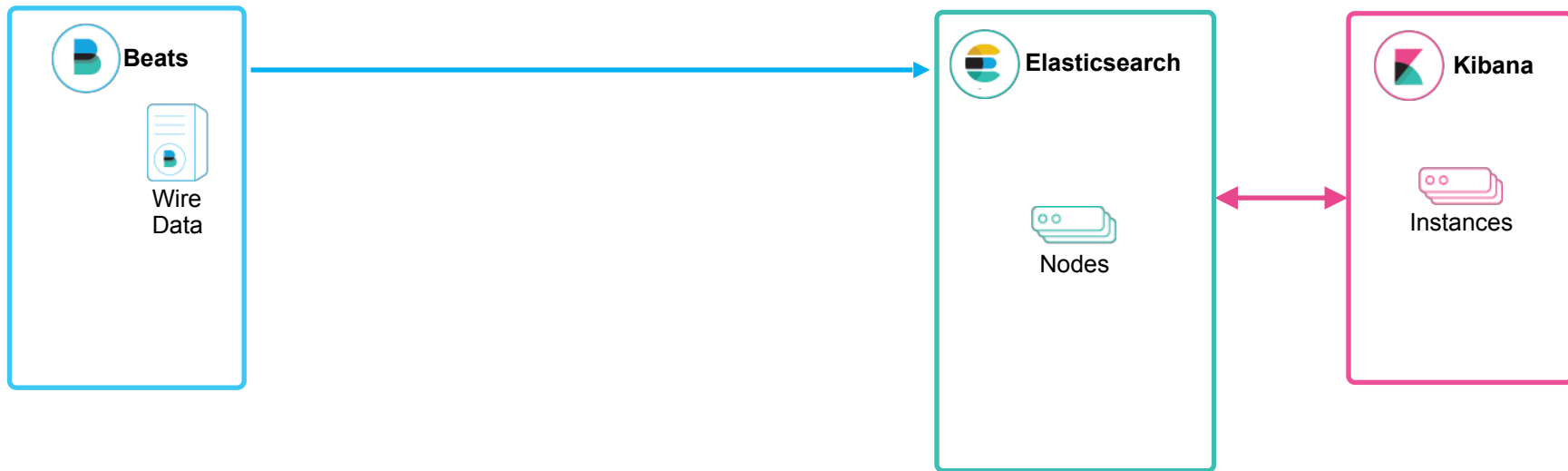
- Metricbeatのsystemモジュールを有効化

```
./metricbeat modules enable system
```

- setupコマンドを実行してからFilebeatを起動

```
./metricbeat setup  
./metricbeat -e
```

PacketbeatでMySQL、PostgreSQLのパケット解析



参考文献

- Elasticsearch - The Definitive guide
 - <http://www.elastic.co/guide/en/elasticsearch/guide/current/index.html>
- 書籍（日本語）
 - データ分析基盤構築入門
 - Elasticsearch実践ガイド



参考サイト

- ユースケース
 - <https://www.elastic.co/use-cases>
- Discuss (Webフォーラム)
 - <https://discuss.elastic.co>
- Elastic{ON}のビデオと資料
 - <https://www.elastic.co/elasticon/videos>
- サポートメニュー
 - <https://www.elastic.co/subscriptions>



Thank you!

- **Web** : <https://www.elastic.co/jp/>
- **Forums** : <https://discuss.elastic.co/>
- **Twitter** : @johtani

