Journée Dev'Ops

Des dashboards pour tous avec ELK

10 Juin 2014
Vincent Spiewak - @vspiewak

Xebia

Introduction

Speaker

Vincent Spiewak
@vspiewak

- 5 ans XP
- Master TA (UPMC)
- http://blog.xebia.fr
- @vspiewak

Agenda

- Introduction
- Logstash
- Monitoring Système
- Monitoring JMX
- Log As A Service
- Monitoring Métier / BI
- Cluster ELK
- Vagrant (démos)

Stack

Logstash

ETL

Elasticsearch

Stockage

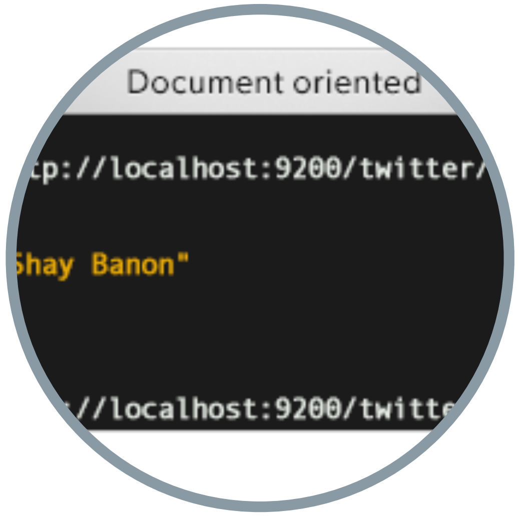
Kibana

Visualisation

Logstash

Inputs	Codecs	Filters	Outputs
41	20	50	55
» stdin	» plain	» grok	» stdout
» file	» json	» date	» file
» udp	» line	» drop	» udp
» tcp	» multiline	» mutate	» tcp
» rabbitmq	» dots	» geoip	» rabbitmq
» s3	» msgpack	» anonymize	» elasticsearch
» …	» …	» …	» …

Elasticsearch



Document oriented

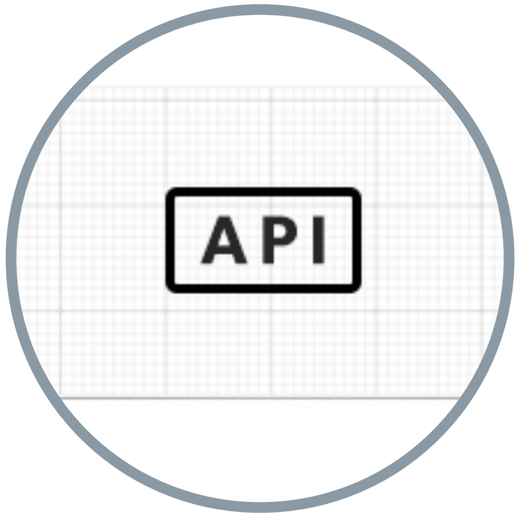
```
tp://localhost:9200/twitter/

Shay Banon"

//localhost:9200/twitte
```

Document

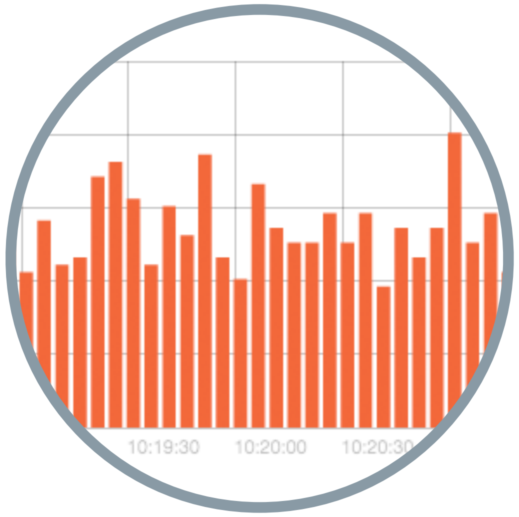
```
url -XPUT ht
'{
  "user": "kim
  "post_date":
  "message":
```

Schema Free

Full Text



REST



Real Time

Distributed

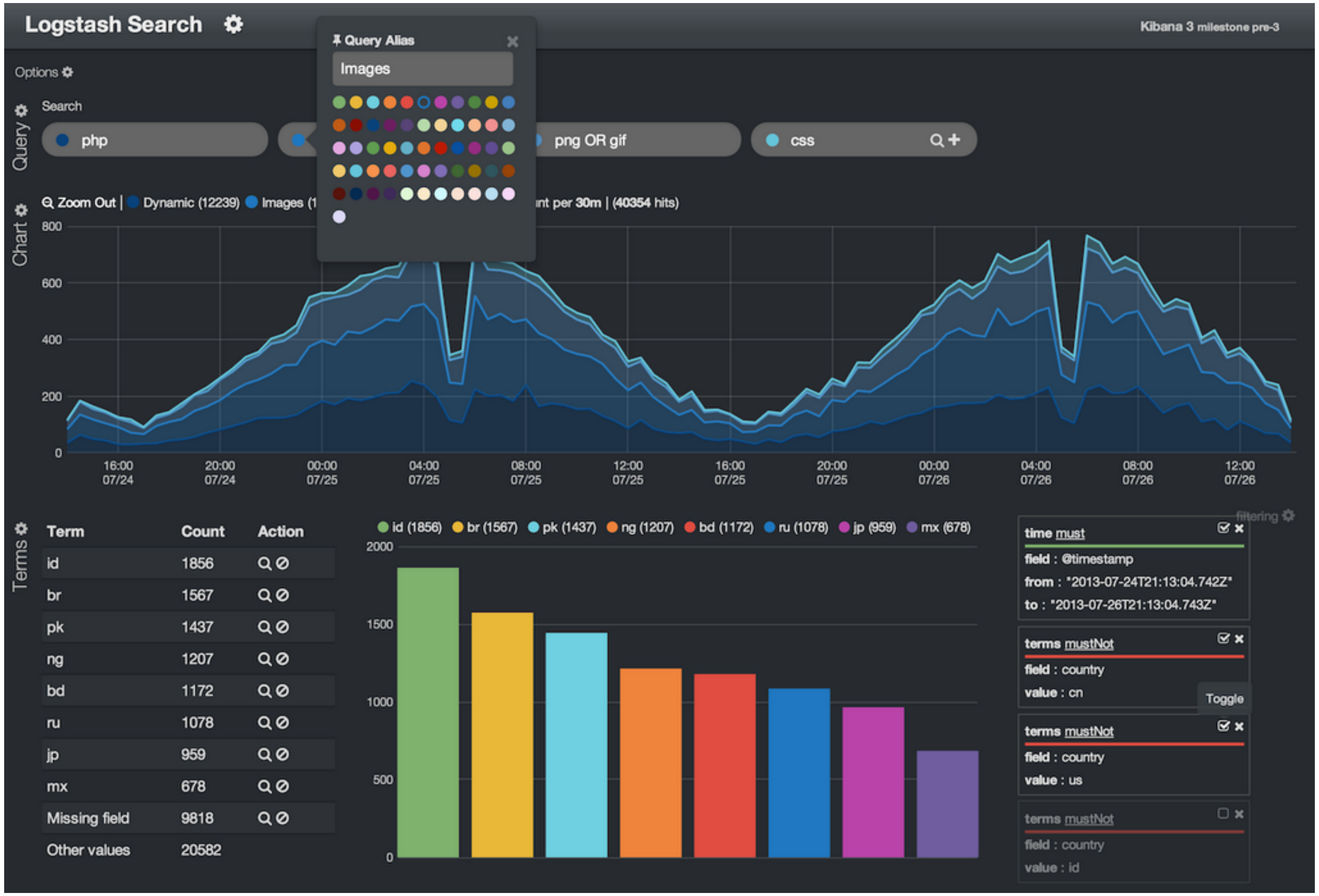
HA

Multi-tenancy

```
p://localhost:9200/kimchy

tp://localhost:9200/elastics

://localhost:9200/elasti
```

Multi–tenancy

Kibana



Term	Count	Action
id	1856	🔍 ⊘
br	1567	🔍 ⊘
pk	1437	🔍 ⊘
ng	1207	🔍 ⊘
bd	1172	🔍 ⊘
ru	1078	🔍 ⊘
jp	959	🔍 ⊘
mx	678	🔍 ⊘
Missing field	9818	🔍 ⊘
Other values	20582	

time must ☑ ✕
field : @timestamp
from : "2013-07-24T21:13:04.742Z"
to : "2013-07-26T21:13:04.743Z"

terms mustNot ☑ ✕
field : country
value : cn

terms mustNot ☑ ✕
field : country
value : us

terms mustNot □ ✕
field : country
value : id

HTML

Dashboards Adaptés

Ops

Infrastructure

- » serveur
- » charge

Dev

Application

- » stacktrace
- » warn, error

Métier

Business

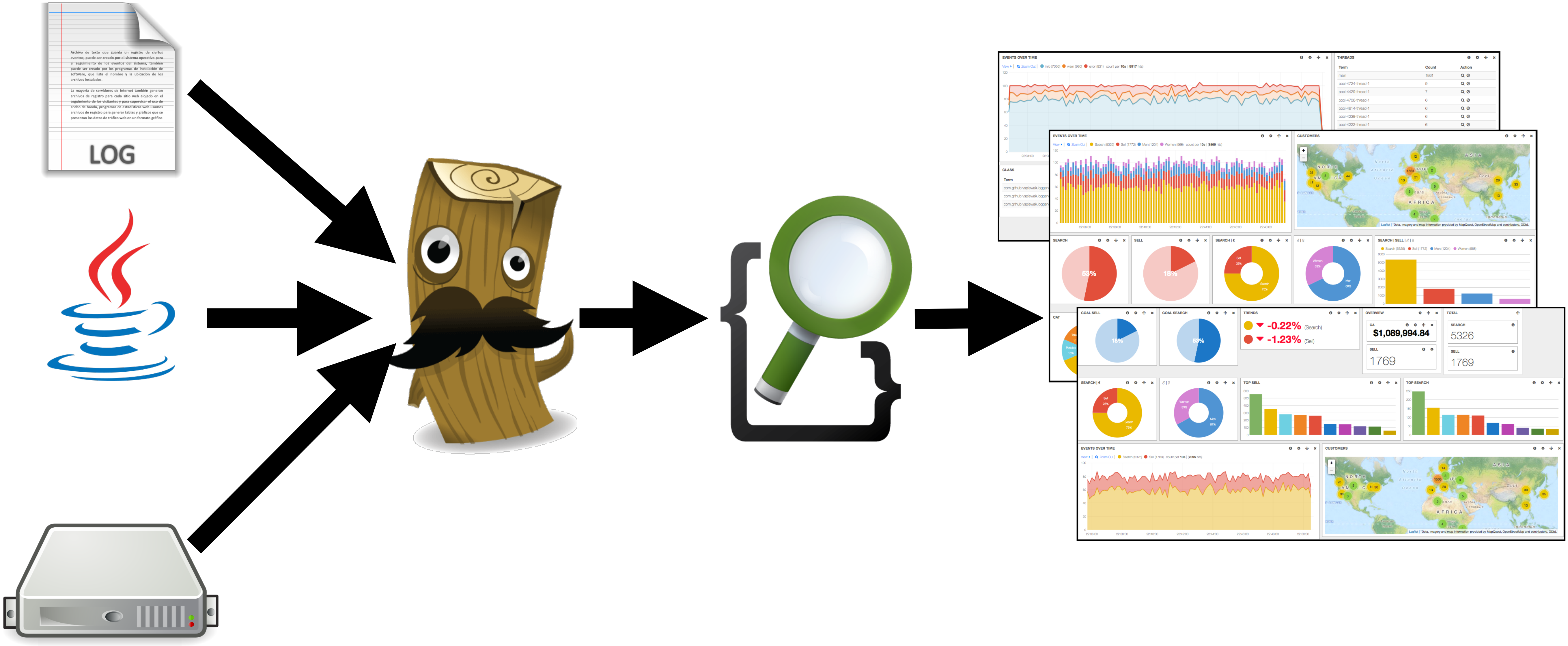
- » client
- » produit

Direction

Objectifs

- » résultat
- » progression

Architecture



Logstash

Logstash – Entrée/Sortie

```
input {
  stdin {}
}

# filters

output {
  stdout { codec => json }
}
```

Logstash – Run

```
$ java -jar logstash.jar agent –f app.conf

2011-04-19T03:44:01.103Z 55.3.244.1 GET /index.html 15824 0.043

{
     "message" => "2011-04-19T03:44:01.103Z GET /index.html 15824 0.043",
  "@timestamp" => "2013-11-03T19:48:53.175Z",
    "@version" => "1",
        "host" => "macbook"
}
```

Logstash – Patterns

https://github.com/logstash/logstash/blob/master/patterns

```
USERNAME [a-zA-Z0-9._-]+

USER %{USERNAME}

INT (?:[+-]?(?:[0-9]+))

WORD \b\w+\b

NOTSPACE \S+

DATA .*?

GREEDYDATA .*

HTTPDATE %{MONTHDAY}/%{MONTH}/%{YEAR}:%{TIME} %{INT}

COMBINEDAPACHELOG %{IPORHOST:clientip} …
```

Logstash – Filtre Grok

```
2011-04-19T03:44:01.103Z 55.3.244.1 GET /index.html 15824 0.043
```

```
filter {
   grok {
      match =>

           [ "message",

             "%{TIMESTAMP_ISO8601:date} %{IP:client} %{WORD:method}
             %{URIPATHPARAM:uri} %{NUMBER:bytes} %{NUMBER:duration}"
           ]
   }
}
```

Filtre Grok – Sortie

```
2011-04-19T03:44:01.103Z 55.3.244.1 GET /index.html 15824 0.043

{
   "@timestamp" => "2013-12-01T21:19:11.303Z",
    "@version" => "1",
       "@bytes" => "15824",
      "@client" => "55.3.244.1",
         "date" => "2011-04-19T03:44:01.103Z",
    "@duration" => "0.043",
         "host" => "macbookpro",
      "message" => "2011-04-19T03:44:01.103Z 55.3.244.1 GET /index.html 15824 0.043",
       "method" => "GET",
          "uri" => "/index.html",
}
```

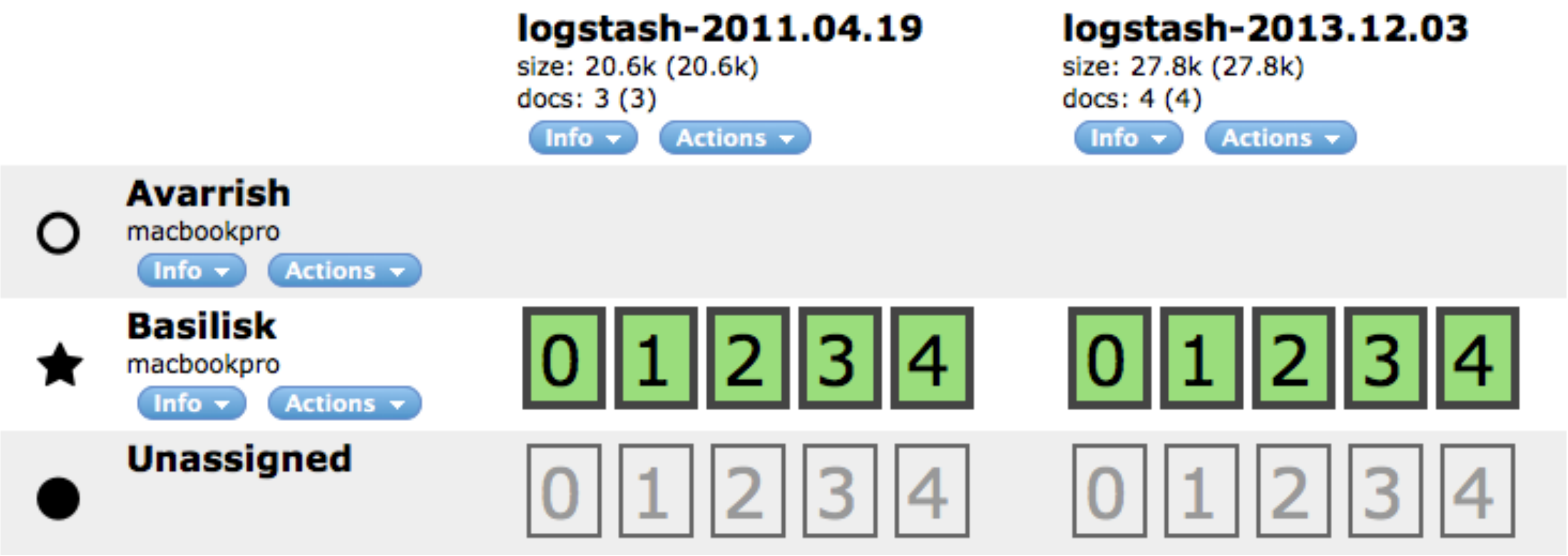
Filtre Date – @Timestamp

```
filter {
    date {
        match => [ "date", "ISO8601" ],
    }
}
```

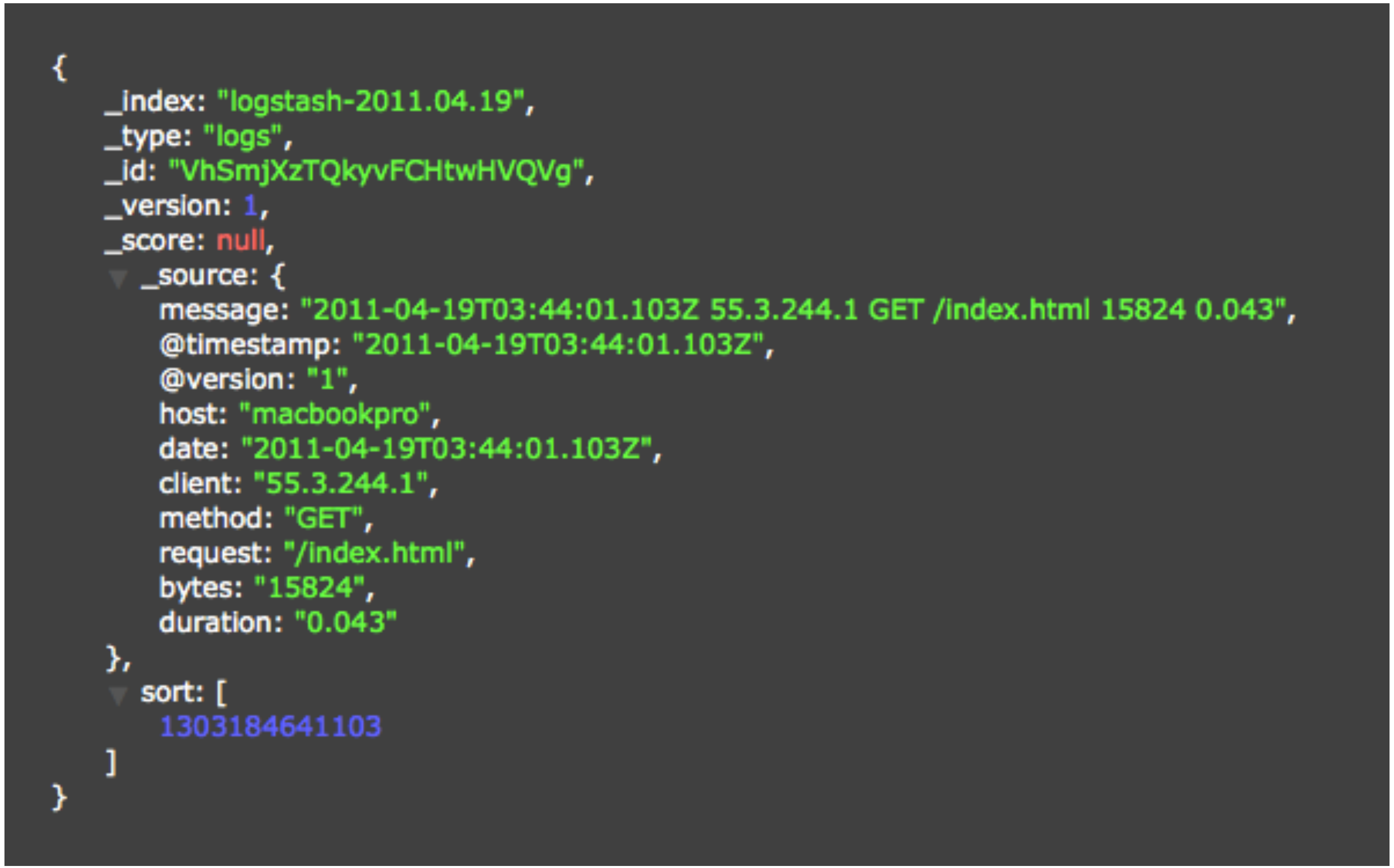
Filtre Date – @Timestamp

```
{
    "@timestamp" => "2011-04-19T03:44:01.103Z",
      "@version" => "1",
        "@bytes" => "15824",
       "@client" => "55.3.244.1",
          "date" => "2011-04-19T03:44:01.103Z",
     "@duration" => "0.043",
          "host" => "macbookpro",
       "message" => "2011-04-19T03:44:01.103Z 55.3.244.1 GET /index.html 15824 0.043",
        "method" => "GET",
           "uri" => "/index.html",
}
```

Sortie Elasticsearch



	logstash-2011.04.19 size: 20.6k (20.6k) docs: 3 (3) Info ▾ Actions ▾	logstash-2013.12.03 size: 27.8k (27.8k) docs: 4 (4) Info ▾ Actions ▾
○ Avarrish macbookpro Info ▾ Actions ▾		
★ Basilisk macbookpro Info ▾ Actions ▾	0 1 2 3 4	0 1 2 3 4
● Unassigned	0 1 2 3 4	0 1 2 3 4



```
{
    _index: "logstash-2011.04.19",
    _type: "logs",
    _id: "VhSmjXzTQkyvFCHtwHVQVg",
    _version: 1,
    _score: null,
    _source: {
       message: "2011-04-19T03:44:01.103Z 55.3.244.1 GET /index.html 15824 0.043",
       @timestamp: "2011-04-19T03:44:01.103Z",
       @version: "1",
       host: "macbookpro",
       date: "2011-04-19T03:44:01.103Z",
       client: "55.3.244.1",
       method: "GET",
       request: "/index.html",
       bytes: "15824",
       duration: "0.043"
    },
    sort: [
       1303184641103
    ]
}
```

Filtres

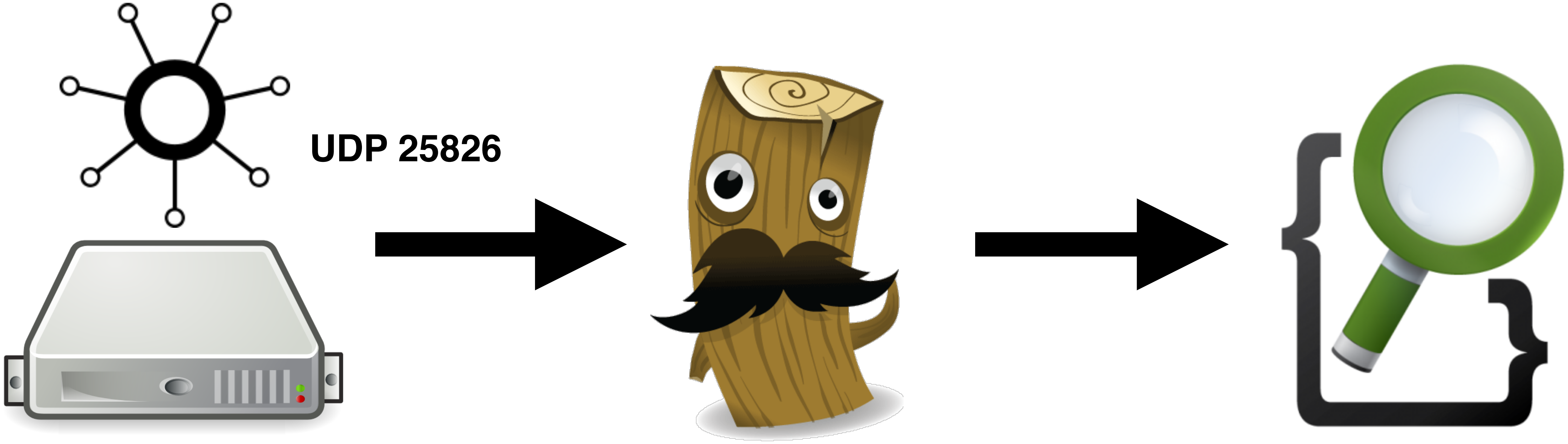
- ajout d'un champ / type / tag
- suppression d'un champ
- split d'un champ
- conversion de type (string, int, float)
- IP => géolocation
- UA => device, browser, os, versions
- conditions
- etc…

Logstash – Sortie Elasticsearch

- host
- port
- cluster
- index => "logstash-%{+YYYY.MM.dd}"
- protocol
- …

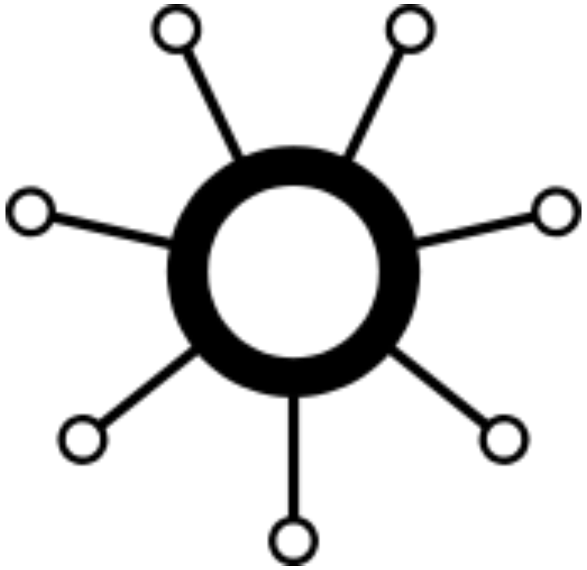
Monitoring Système

Monitoring Système: Collectd



Collectd – Plugins

- cpu (jiffies)
- process
- users
- disk
- memory
- swap
- network
- Java / JMX
- MySQL
- …



Collectd – Configuration

```
FQDNLookup true
LoadPlugin syslog
<Plugin syslog>
 LogLevel info
</Plugin>
LoadPlugin cpu
LoadPlugin df
LoadPlugin disk
LoadPlugin entropy
LoadPlugin interface
LoadPlugin irq
LoadPlugin load
LoadPlugin memory
LoadPlugin network
LoadPlugin processes
LoadPlugin rrdtool
LoadPlugin swap
LoadPlugin users
<Plugin interface>
 Interface "eth0"
 IgnoreSelected false
</Plugin>
<Plugin network>
    <Server "127.0.0.1" "25826">
    </Server>
</Plugin>
<Plugin rrdtool>
 DataDir "/var/lib/collectd/rrd"
</Plugin>
Include "/etc/collectd/filters.conf"
Include "/etc/collectd/thresholds.conf"
```

Collectd – Logstash Conf

```
input {
  collectd {
     host => "127.0.0.1"
  }
}

output {
    elasticsearch {}
}
```

Elasticsearch – Samples

```
{
    "@version": "1",
    "@timestamp": "2014-06-09T23:01:11.000Z",
    "host": "precise64",
    "plugin": "memory",
    "collectd_type": "memory",
    "type_instance": "cached",
    "value": 267845632
}

{
    "@version": "1",
    "@timestamp": "2014-06-09T23:01:11.000Z",
    "host": "precise64",
    "plugin": "memory",
    "collectd_type": "memory",
    "type_instance": "used",
    "value": 703348736
}
```

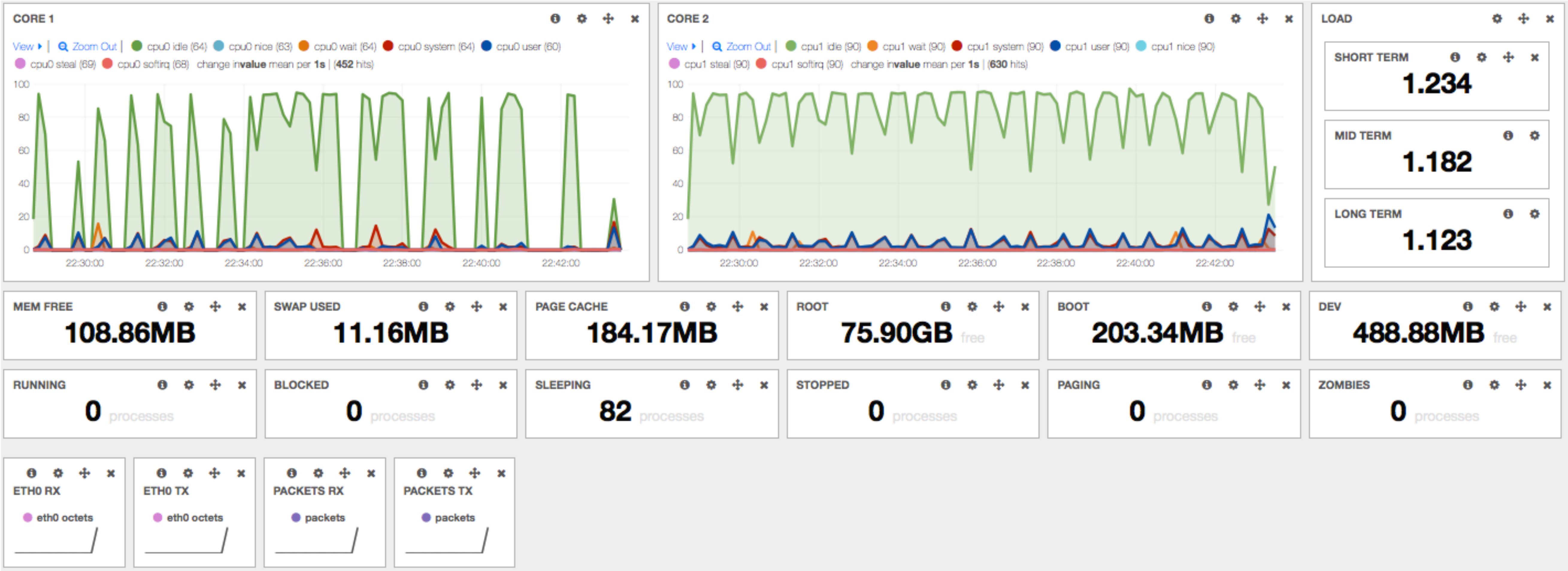
Monitoring Système

system-survey

Démo

Monitoring Système

system-survey



CORE 1

View ▸ | Zoom Out | ● cpu0 idle (64) ● cpu0 nice (63) ● cpu0 wait (64) ● cpu0 system (64) ● cpu0 user (60) ● cpu0 steal (69) ● cpu0 softirq (68) change in**value** mean per **1s** | (**452** hits)

CORE 2

View ▸ | Zoom Out | ● cpu1 idle (90) ● cpu1 wait (90) ● cpu1 system (90) ● cpu1 user (90) ● cpu1 nice (90) ● cpu1 steal (90) ● cpu1 softirq (90) change in**value** mean per **1s** | (**630** hits)

LOAD

SHORT TERM 1.234

MID TERM 1.182

LONG TERM 1.123

MEM FREE 108.86MB

SWAP USED 11.16MB

PAGE CACHE 184.17MB

ROOT 75.90GB free

BOOT 203.34MB free

DEV 488.88MB free

RUNNING 0 processes

BLOCKED 0 processes

SLEEPING 82 processes

STOPPED 0 processes

PAGING 0 processes

ZOMBIES 0 processes

ETH0 RX

● eth0 octets

ETH0 TX

● eth0 octets

PACKETS RX

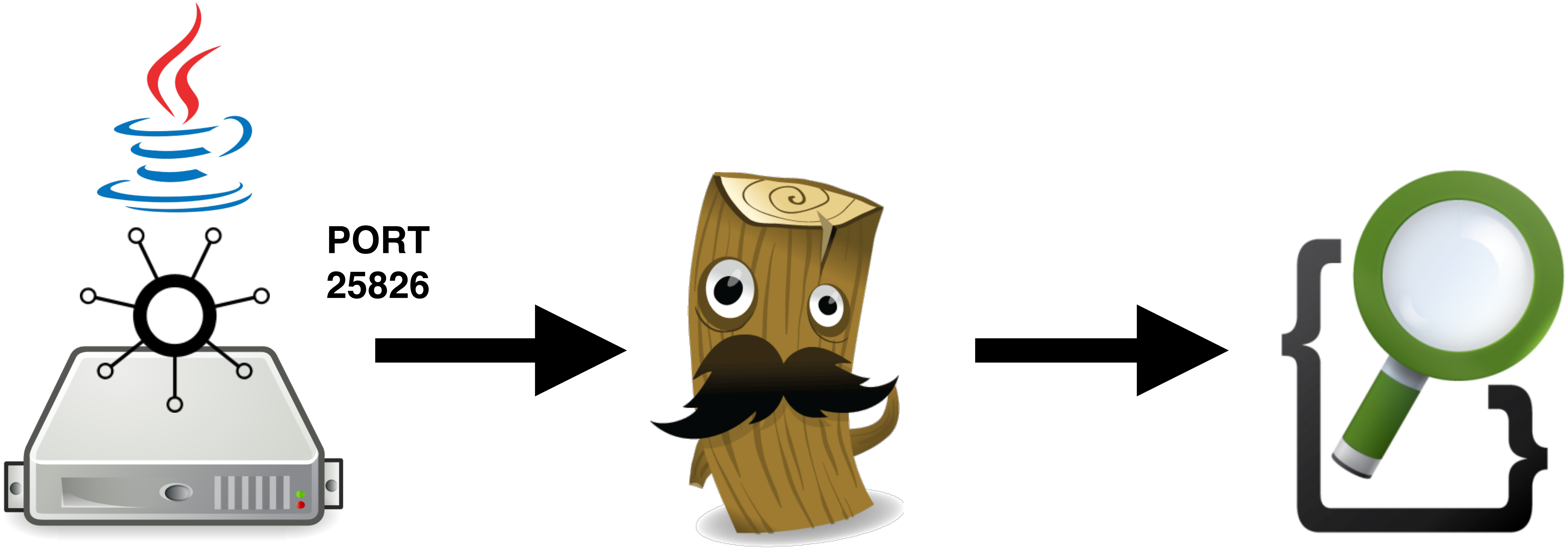
● packets

PACKETS TX

● packets

Monitoring JMX

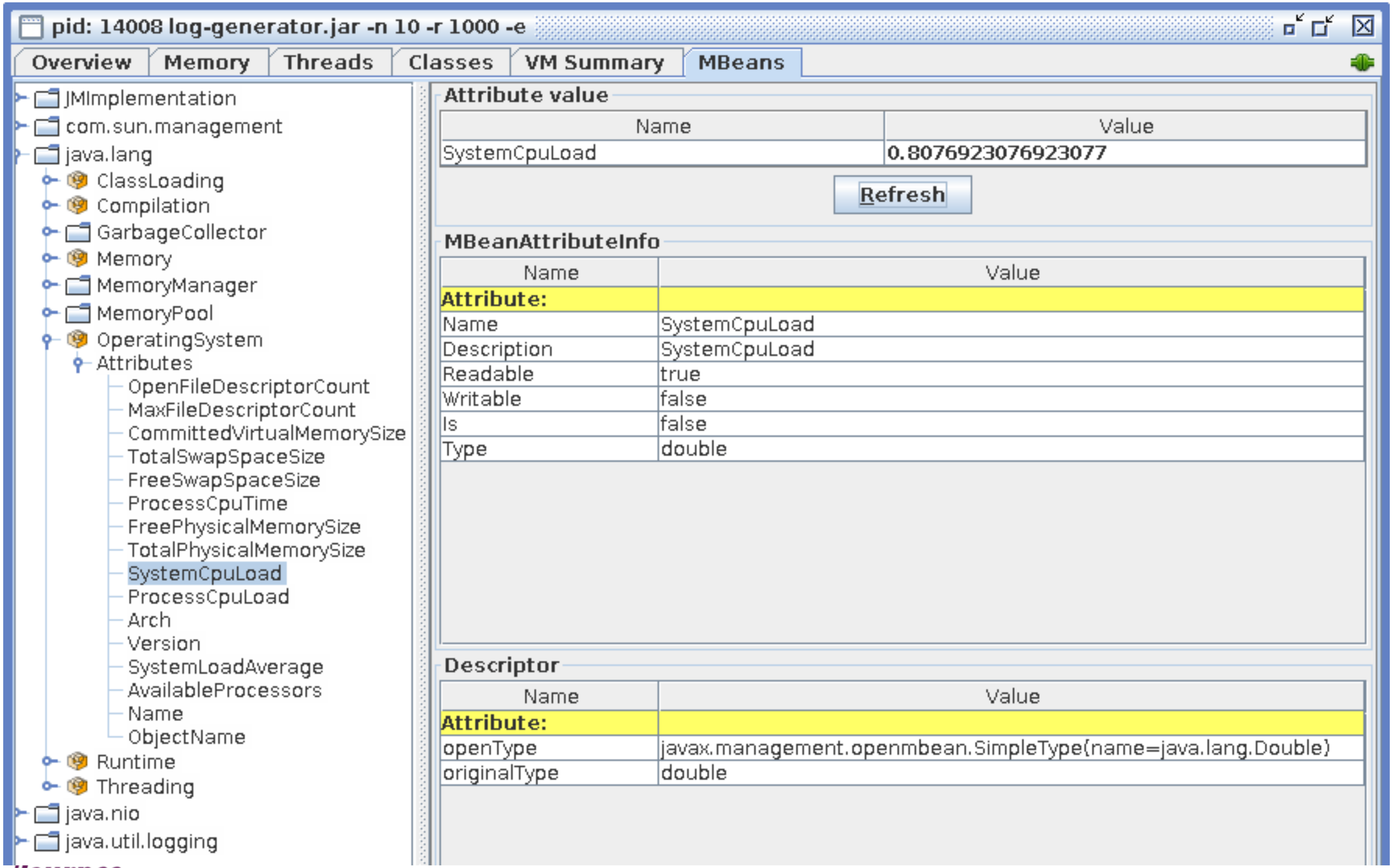
Monitoring JMX: Collectd JMX



Collectd – Setup Java & JMX

```
# check dynamic libraries

ldd /usr/lib/collectd/java.so



# fix libjvm.so not found error

ln -s /usr/lib/jvm/java-7-openjdk-amd64/jre/lib/amd64/server/libjvm.so /usr/lib/libjvm.so
```

JConsole – SystemCpuLoad



Attribute value

Name	Value
SystemCpuLoad	0.8076923076923077

MBeanAttributeInfo

Name	Value
Attribute:	
Name	SystemCpuLoad
Description	SystemCpuLoad
Readable	true
Writable	false
Is	false
Type	double

Descriptor

Name	Value
Attribute:	
openType	javax.management.openmbean.SimpleType(name=java.lang.Double)
originalType	double

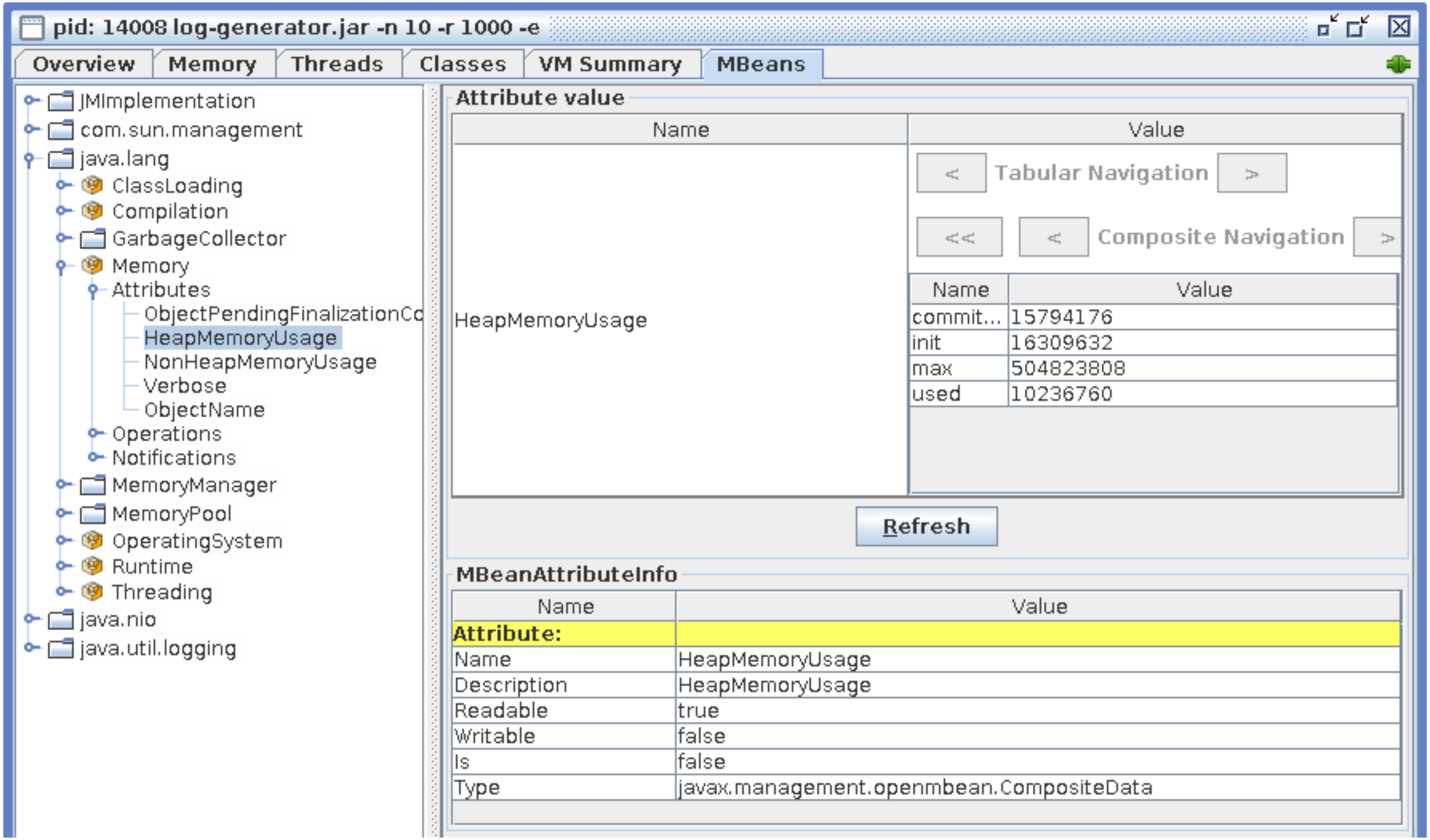
Collectd – Plugin Java & JMX

```
<Plugin "java">
 JVMARG "-Djava.class.path=/usr/share/collectd/java/collectd-api.jar:/usr/share/collectd/java/generic-jmx.jar"
  LoadPlugin "org.collectd.java.GenericJMX"
  <Plugin "GenericJMX">

   <MBean "os">
      ObjectName "java.lang:type=OperatingSystem"

      <Value>
        Type "gauge"
        InstancePrefix "system_cpu_load"
        Attribute "SystemCpuLoad"
      </Value>

    </MBean>

    <Connection>
      ServiceURL "service:jmx:rmi:///jndi/rmi://localhost:9010/jmxrmi"
      Collect "os"
   </Connection>

  </Plugin>
</Plugin>
```

Elasticsearch – Samples

```
{
    "@version": "1",
    "@timestamp": "2014-06-09T23:01:11.000Z",
    "host": "localhost",
    "plugin": "GenericJMX",
    "collectd_type": "gauge",
    "type_instance": "system_cpu_load",
    "value": 0.5587837837837838
}
```

JConsole – HeapMemoryUsage



Attribute value

Name	Value
HeapMemoryUsage	

Name	Value
commit...	15794176
init	16309632
max	504823808
used	10236760

MBeanAttributeInfo

Name	Value
Attribute:	
Name	HeapMemoryUsage
Description	HeapMemoryUsage
Readable	true
Writable	false
Is	false
Type	javax.management.openmbean.CompositeData

Collectd – JMX – Type Table

```
# Heap memory usage
<MBean "memory-heap">
  ObjectName "java.lang:type=Memory"
  #InstanceFrom ""
  InstancePrefix "memory-heap"

  # Creates four values: committed, init, max, used
  <Value>
    Type "jmx_memory"
    Table true
    Attribute "HeapMemoryUsage"
  </Value>
</MBean>
```

Collectd – Types

```
/usr/share/collectd/types.db

gauge                   value:GAUGE:U:U
load                    shortterm:GAUGE:0:100, midterm:GAUGE:0:100, longterm:GAUGE:0:100
percent                 percent:GAUGE:0:100.1
jmx_memory              value:GAUGE:0:U
```

Elasticsearch – Samples

```
{
    "@version": "1",
    "@timestamp": "2014-06-09T23:01:11.000Z",
    "host": "localhost",
    "plugin": "GenericJMX",
    "plugin_instance": "memory-heap",
    "collectd_type": "jmx_memory",
    "type_instance": "used",
    "value": 62282808
}

{
    "@version": "1",
    "@timestamp": "2014-06-09T23:01:11.000Z",
    "host": "localhost",
    "plugin": "GenericJMX",
    "plugin_instance": "memory-heap",
    "collectd_type": "jmx_memory",
    "type_instance": "init",
    "value": 104857600
}
```

Collectd – Custom MBean

```
<MBean "flume-source">
  ObjectName "org.apache.flume.source:type=source-1"
  InstancePrefix "flume-source-1"
 <Value>
    Type "gauge"
    InstancePrefix "event_received_count"
    Attribute "EventReceivedCount"
  </Value>
  <Value>
    Type "gauge"
    InstancePrefix "event_accepted_count"
    Table false
    Attribute "EventAcceptedCount"
  </Value>
</MBean>
```

Elasticsearch – Samples

```
{
    "@version": "1",
    "@timestamp": "2014-06-09T23:09:41.000Z",
    "host": "localhost",
    "plugin": "GenericJMX",
    "plugin_instance": "flume-source-1",
    "collectd_type": "gauge",
    "type_instance": "event_accepted_count",
    "value": 1246501
}
```

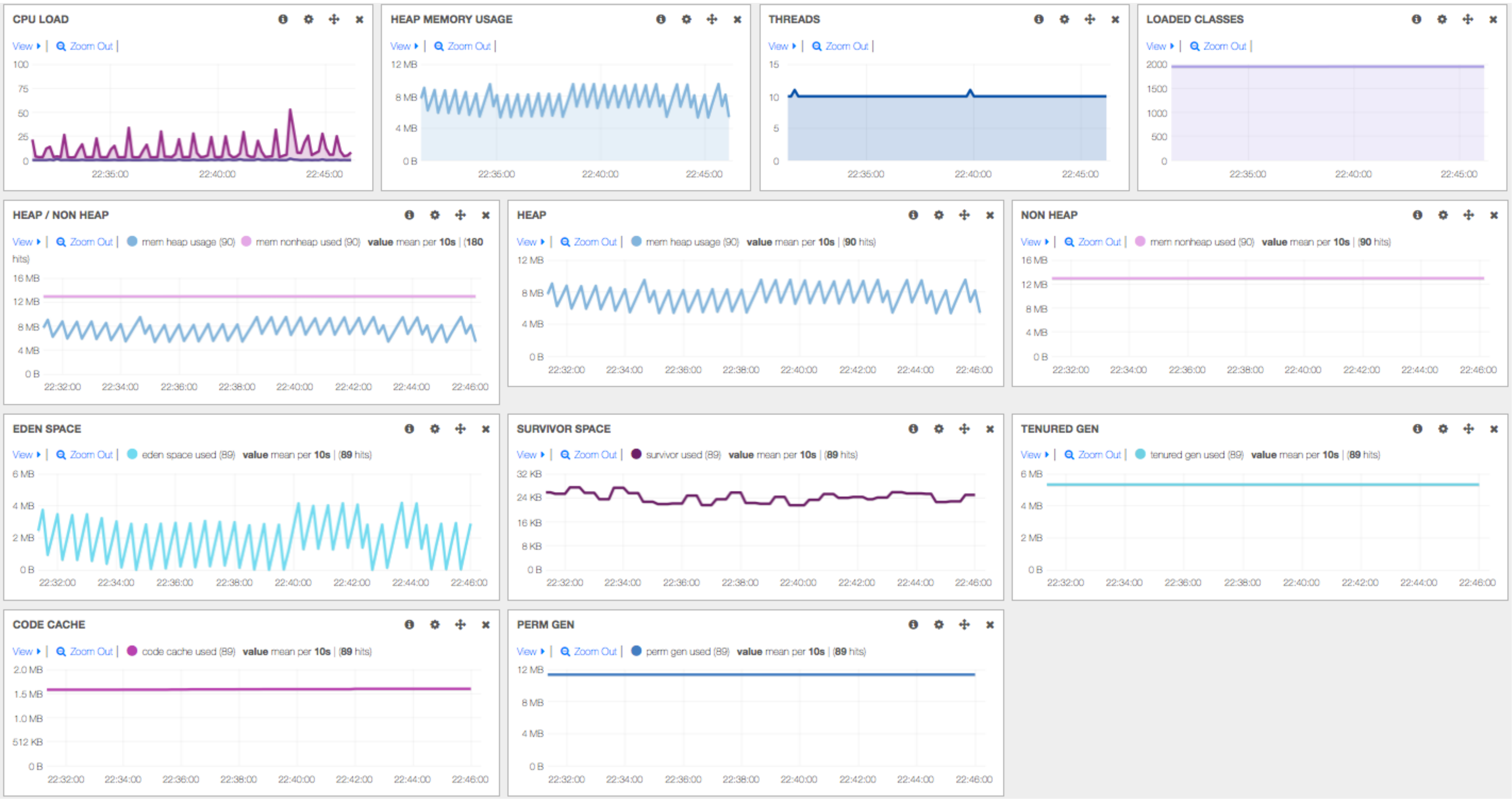
Monitoring JVM / JMX

Flume JMX

Démo

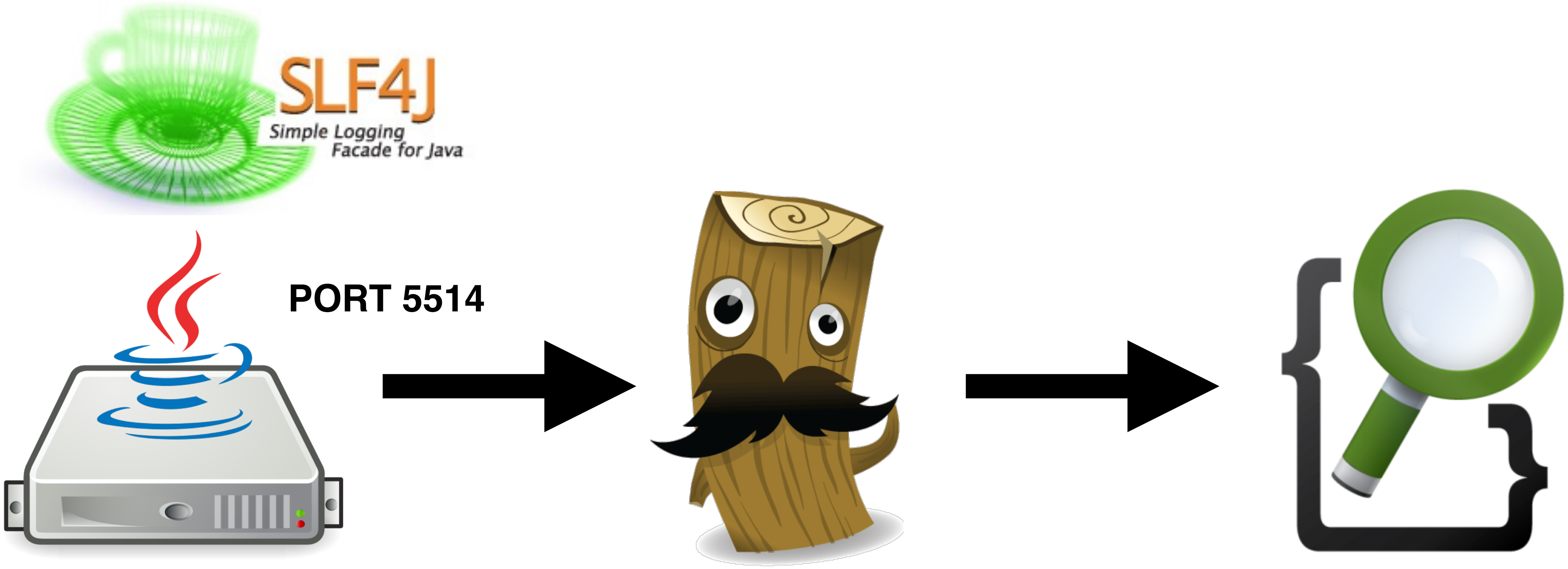
Monitoring JVM / JMX

Flume JMX



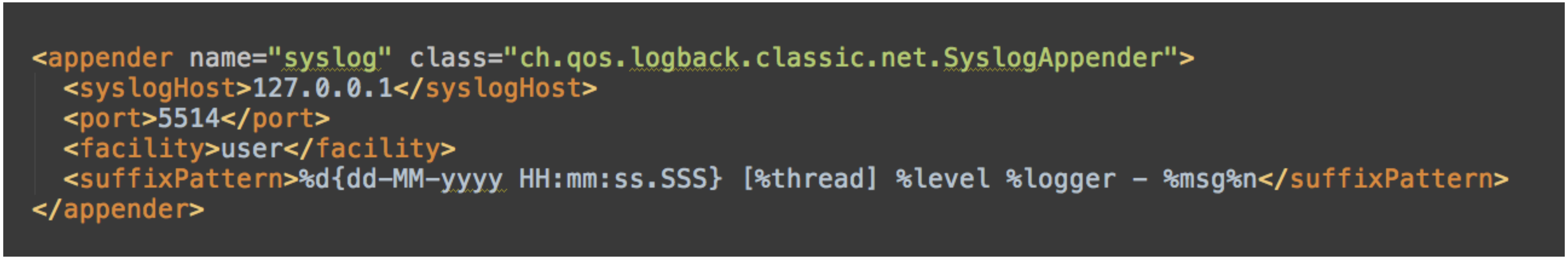
Log As Service

Log As A Service: SyslogAppender



Log As A Service – Logback

SyslogAppender



```xml
<appender name="syslog" class="ch.qos.logback.classic.net.SyslogAppender">
  <syslogHost>127.0.0.1</syslogHost>
  <port>5514</port>
  <facility>user</facility>
  <suffixPattern>%d{dd-MM-yyyy HH:mm:ss.SSS} [%thread] %level %logger - %msg%n</suffixPattern>
</appender>
```

Logstash – Syslog configuration

```
input {
  udp {
     port => "5514"
  }
}

filter {
   grok {
      patterns_dir => "./patterns"
      match => ["message","%{LOGBACK_SYSLOG}"]
   }
}

filter {
   date {
      match => ["log_date","dd-MM-YYYY HH:mm:ss.SSS"]
   }
}

output {
  elasticsearch {}
}
```

Logstash – Pattern

```
LOG_DATE %{MONTHDAY}-%{MONTHNUM}-%{YEAR} %{HOUR}:%{MINUTE}:%{SECOND}.[0-9]{3}

SYSLOG_BASE %{SYSLOG5424PRI}%{SYSLOGTIMESTAMP:syslog_timestamp} %{SYSLOGHOST:syslog_host}
SYSLOG %{SYSLOG_BASE} %{GREEDYDATA:syslog_message}

LOGBACK_SYSLOG_BASE %{SYSLOG_BASE} %{LOG_DATE:log_date} \[%{NOTSPACE:thread}\] %{LOGLEVEL:log_level} %{NOTSPACE:classname}
LOGBACK_SYSLOG %{LOGBACK_SYSLOG_BASE} %{GREEDYDATA:log_msg}
```

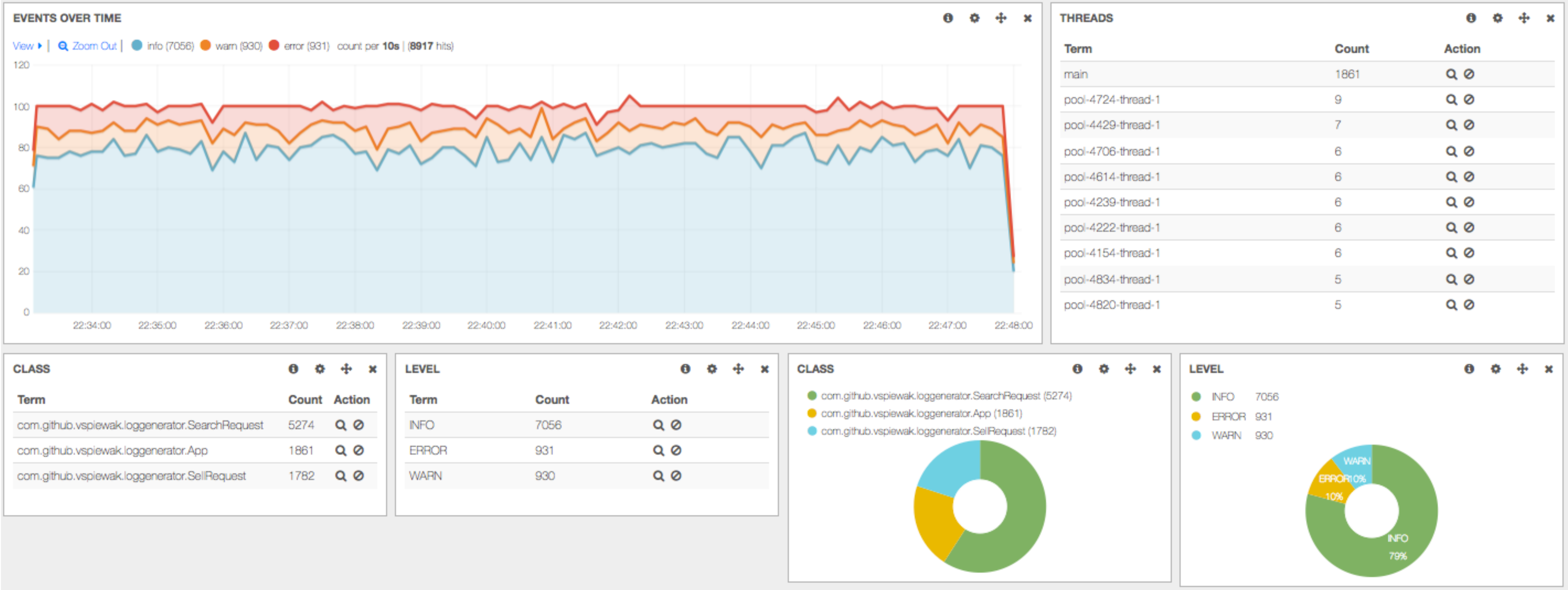
Monitoring Log

Syslog

Démo

Monitoring Log

Syslog



EVENTS OVER TIME

View ▸ | 🔍 Zoom Out | ● info (7056) ● warn (930) ● error (931) count per **10s** | (**8917** hits)

THREADS

Term	Count	Action
main	1861	🔍 ⊘
pool-4724-thread-1	9	🔍 ⊘
pool-4429-thread-1	7	🔍 ⊘
pool-4706-thread-1	6	🔍 ⊘
pool-4614-thread-1	6	🔍 ⊘
pool-4239-thread-1	6	🔍 ⊘
pool-4222-thread-1	6	🔍 ⊘
pool-4154-thread-1	6	🔍 ⊘
pool-4834-thread-1	5	🔍 ⊘
pool-4820-thread-1	5	🔍 ⊘

CLASS

Term	Count	Action
com.github.vspiewak.loggenerator.SearchRequest	5274	🔍 ⊘
com.github.vspiewak.loggenerator.App	1861	🔍 ⊘
com.github.vspiewak.loggenerator.SellRequest	1782	🔍 ⊘

LEVEL

Term	Count	Action
INFO	7056	🔍 ⊘
ERROR	931	🔍 ⊘
WARN	930	🔍 ⊘

CLASS

- ● com.github.vspiewak.loggenerator.SearchRequest (5274)
- ● com.github.vspiewak.loggenerator.App (1861)
- ● com.github.vspiewak.loggenerator.SellRequest (1782)

LEVEL

●	INFO	7056
●	ERROR	931
●	WARN	930

Monitoring Métier– BI

GeekShop

Problème

- Quels sont les produits les plus achetés ?
- Quelle est la répartition H/F de mes clients ?
- Quels sont mes clients les plus fidèles ?
- Combien de femmes à Paris ont acheté un iPod Touch Bleu 32 Go entre le 12 octobre 2012 à 14h30 et le 4 novembre 2013 à 19h ?

GeekShop – Format Logs

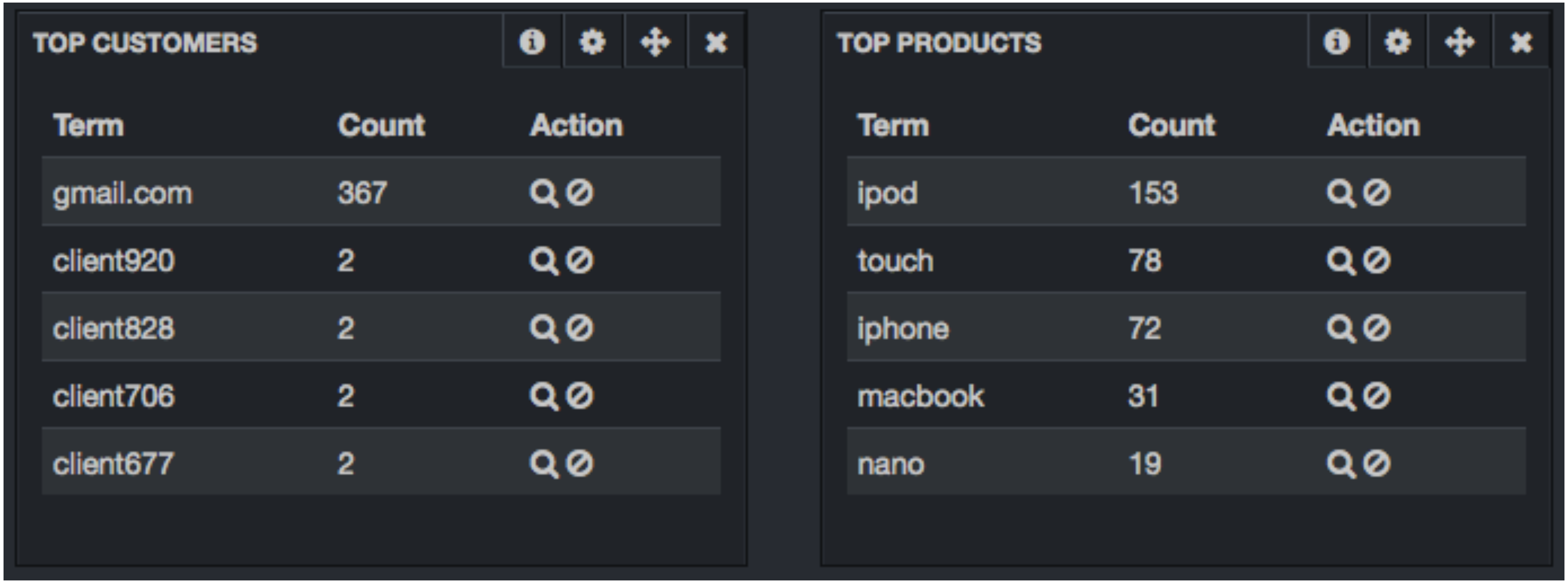
```
09-06-2014 21:27:42.228 [pool-32-thread-1] INFO
com.github.vspiewak.loggenerator.SearchRequest -
id=317&ua=Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/534.24
(KHTML, like Gecko) Chrome/11.0.696.65 Safari/
534.24&ip=94.228.34.210&category=Mobile
```

```
09-06-2014 21:27:42.227 [pool-32-thread-1] INFO
com.github.vspiewak.loggenerator.SellRequest - id=313&ua=Mozilla/
5.0 (Windows; U; Windows NT 6.1; en-US; rv:1.9.2.17) Gecko/
20110420 Firefox/
3.6.17&ip=202.46.52.35&email=client314@gmail.com&sex=M&brand=Appl
e&name=iPod Touch&model=iPod Touch - Jaune - Disque
32Go&category=Baladeur&color=Jaune&options=Disque
32Go&price=329.0
```

GeekShop – Après Logstash

```
{
    "_index": "logstash-2014.06.09",
    "_type": "app-log",
    "_id": "gaQXRn9mROiAGjhBZ2h2Og",
    "_version": 1,
    "found": true,
    "_source": {
        "message": "09-06-2014 21:27:42.228 [pool-32-thread-1] INFO com.github.vspiewak.loggenerator.SearchRequest - id=317&ua=Mozilla/5.0 (X11;
Linux x86_64) AppleWebKit/534.24 (KHTML, like Gecko) Chrome/11.0.696.65 Safari/534.24&ip=94.228.34.210&category=Mobile",
        "@version": "1",
        "@timestamp": "2014-06-09T19:27:42.228Z",
        "type": "app-log",
        "host": "precise64",
        "path": "/home/vagrant/app.log",
        "log_date": "09-06-2014 21:27:42.228",
        "thread": "pool-32-thread-1",
        "log_level": "INFO",
        "classname": "com.github.vspiewak.loggenerator.SearchRequest",
        "log_msg": "- id=317&ua=Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/534.24 (KHTML, like Gecko) Chrome/11.0.696.65 Safari/
534.24&ip=94.228.34.210&category=Mobile",
        "id": 317,
        "ua": "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/534.24 (KHTML, like Gecko) Chrome/11.0.696.65 Safari/534.24",
        "ip": "94.228.34.210",
        "category": "Mobile",
        "tags": [
            "search"
        ],
        "geoip": {
            "ip": "94.228.34.210",
            "country_code2": "GB",
            "country_code3": "GBR",
            "country_name": "United Kingdom",
            "continent_code": "EU",
            "latitude": 51.5,
            "longitude": -0.12999999999999545,
            "timezone": "Europe/London",
            "location": [
                -0.12999999999999545,
                51.5
            ]
        },
        "useragent": {
            "name": "Chrome",
            "os": "Linux",
            "os_name": "Linux",
            "device": "Other",
            "major": "11",
            "minor": "0",
            "patch": "696"
        }
    }
}
```

Kibana: Terms & Analysers



TOP CUSTOMERS

Term	Count	Action
gmail.com	367	🔍 ⊘
client920	2	🔍 ⊘
client828	2	🔍 ⊘
client706	2	🔍 ⊘
client677	2	🔍 ⊘

TOP PRODUCTS

Term	Count	Action
ipod	153	🔍 ⊘
touch	78	🔍 ⊘
iphone	72	🔍 ⊘
macbook	31	🔍 ⊘
nano	19	🔍 ⊘

Elasticsearch Template Mapping

Change analyser on specific indexes & fields

```
curl -XPUT http://localhost:9200/_template/logstash_per_index -d '{
    "template" : "logstash*",
    "mappings" : {
        "_default_" : {
           "properties" : {
                "@timestamp": { "type": "date", "index": "not_analyzed" },
                "ip": { "type" : "ip", "index": "not_analyzed" },
                "name": { "type" : "string", "index": "not_analyzed" },
                "options": { "type" : "string", "index": "not_analyzed" },
                "email": { "type" : "string", "index": "not_analyzed" }
            }
        }
    }
}'
```

Journée DEV;OPS

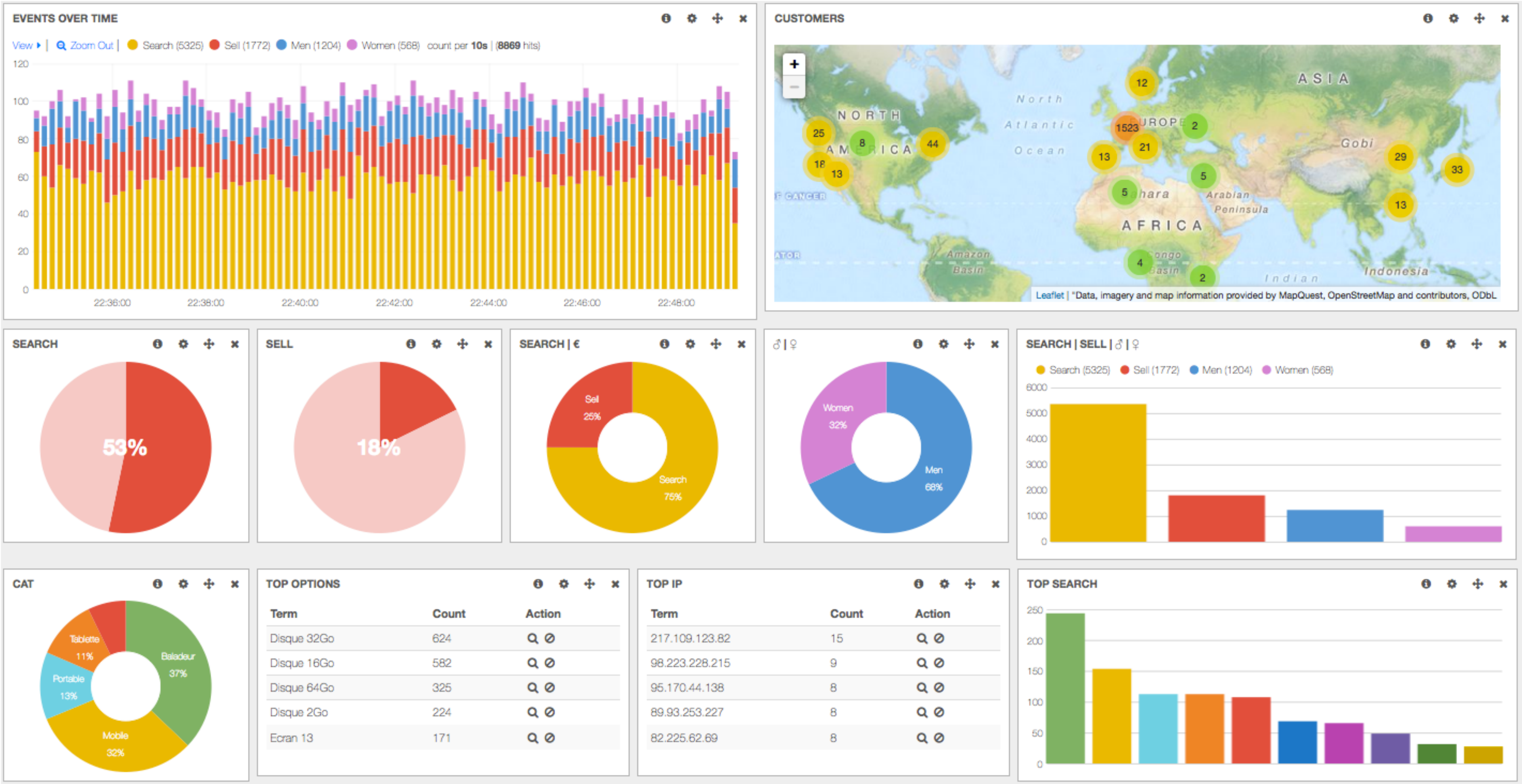
Monitoring Métier / Business

eshop-survey

Démo

Monitoring Métier / Business

eshop-survey



TOP OPTIONS

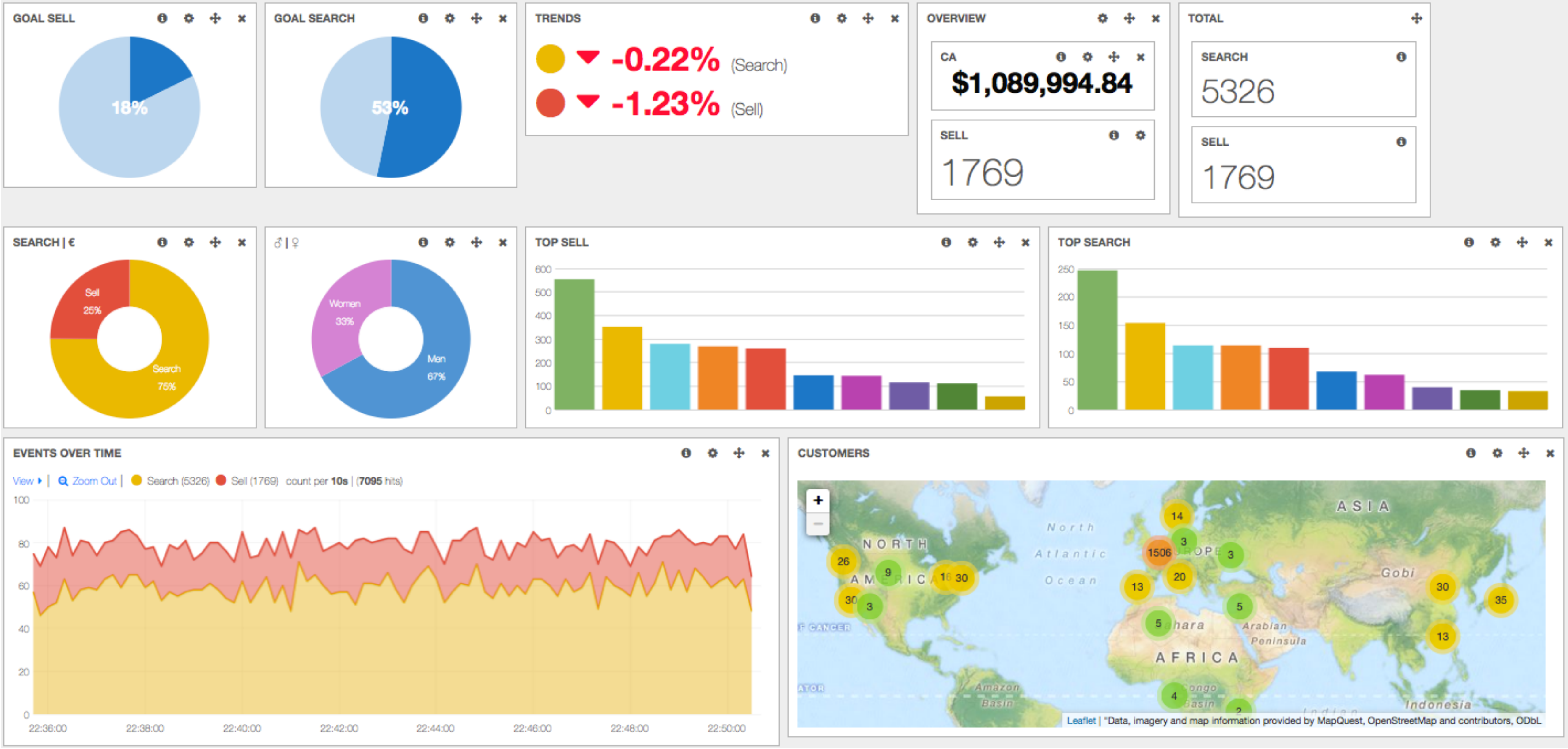
Term	Count	Action
Disque 32Go	624	
Disque 16Go	582	
Disque 64Go	325	
Disque 2Go	224	
Ecran 13	171	

TOP IP

Term	Count	Action
217.109.123.82	15	
98.223.228.215	9	
95.170.44.138	8	
89.93.253.227	8	
82.225.62.69	8	

Monitoring Métier / Business

eshop-survey



● ▼ -0.22% (Search)

● ▼ -1.23% (Sell)

OVERVIEW

CA $1,089,994.84

SELL 1769

TOTAL

SEARCH 5326

SELL 1769

Cluster ELK

Elasticsearch – Feedbacks

- The Guardian: social network – real time feedback
- StackOverflow: full-text search with geolocation and « more like »
- Goldman Sacks: 5TB logs/day + analysis stock market
- …

Elasticsearch – NoSQL

SQL	Partitions	DB	Table	Ligne	Colonne
ES	Cluster	Indices	Type	Document	Champ

Elasticsearch – Types de noeuds

- master
- data
- search

Elasticsearch – Shard & Replica

- shards → +indexing, +distribution (one-time setting)

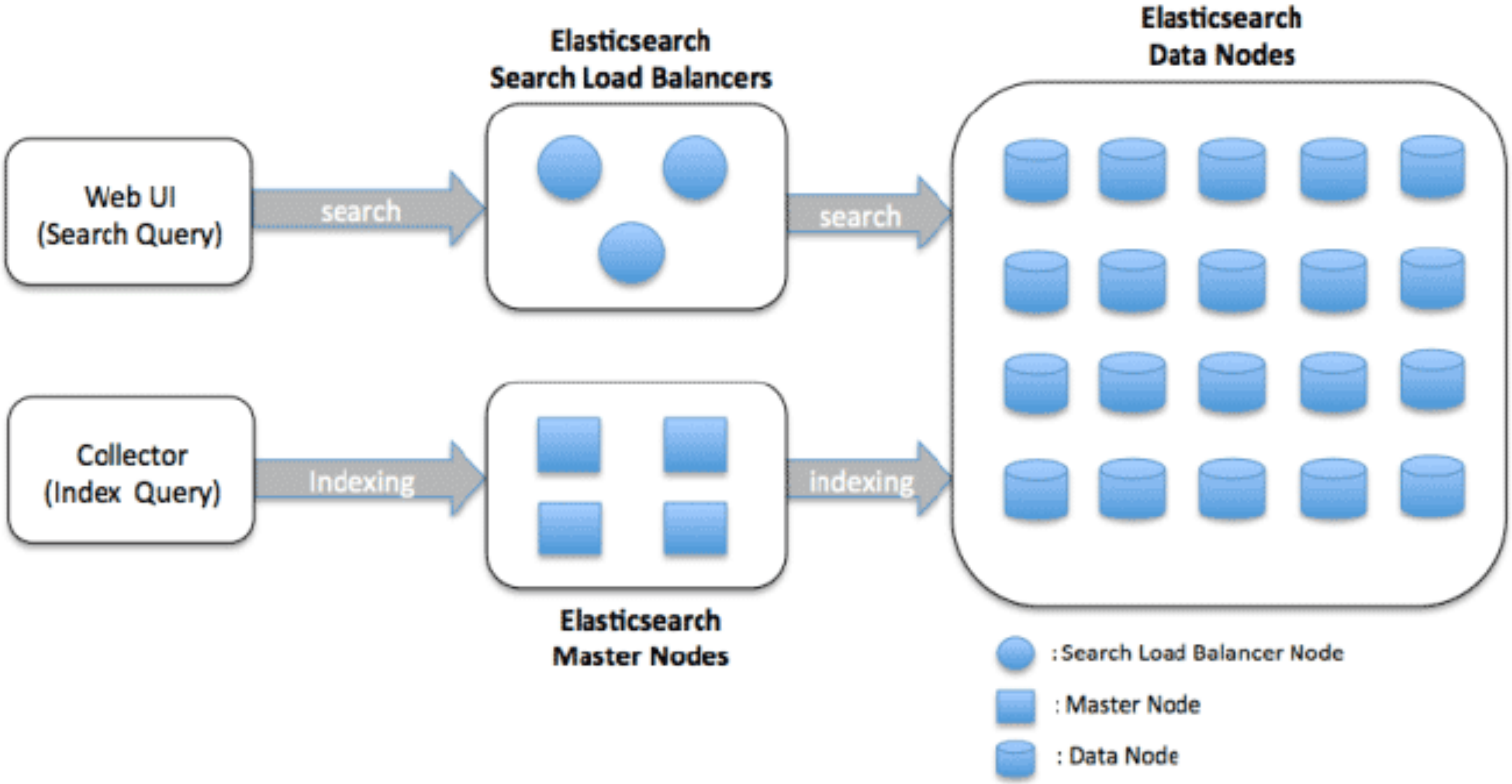
- replicas → +search, +availability

Elasticsearch – Health

- GREEN → all primaries/replicas shards active
- YELLOW → all primaries shards active
- RED → not all primaries shards

Cluster Elasticsearch

http://www.cubrid.org/blog/dev-platform/our-experience-creating-large-scale-log-search-system-using-elasticsearch/



Cluster Elasticsearch

es-cluster

Démo

Vagrant (Démos)

Démo

Pré-requis

Virtual Box

Vagrant

Git *

Démo @ Home

- https://github.com/vspiewak/elk-devops-day-2014

```
$ tree -L 1 .
.
├── README.md
├── demo-all
├── es-cluster
├── eshop-survey
├── flume-jmx
├── slides
├── syslog
└── system-survey
```

Vagrant – Shortcuts

- cd demo-all

- vagrant up
- vagrant ssh
- sudo jconsole
- vagrant halt*
- vagrant destroy

Vagrant VM

- config.vm.box = "hashicorp/precise64"

- config.vm.network "forwarded_port", guest: 80, host: 10080
- config.vm.network "forwarded_port", guest: 9200, host: 19200

- config.ssh.forward_x11 = true

- vb.customize ["modifyvm", :id, "––ioapic", "on", "––cpuexecutioncap", "40", "––cpus", "2", "––memory", "1024"]

- bootstrap.sh

Questions ?

THANK YOU

FOR watching