ELASTICSEARCH | LOGSTASH | KIBANA

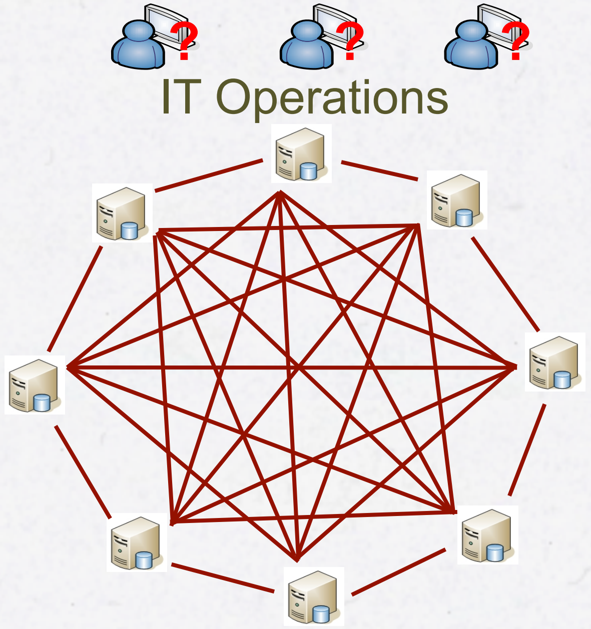
HANS THUNBERG
OLA DEIBITSCH

2015-01-28 | CALLISTAENTERPRISE.SE

CALLISTA
— ENTERPRISE —

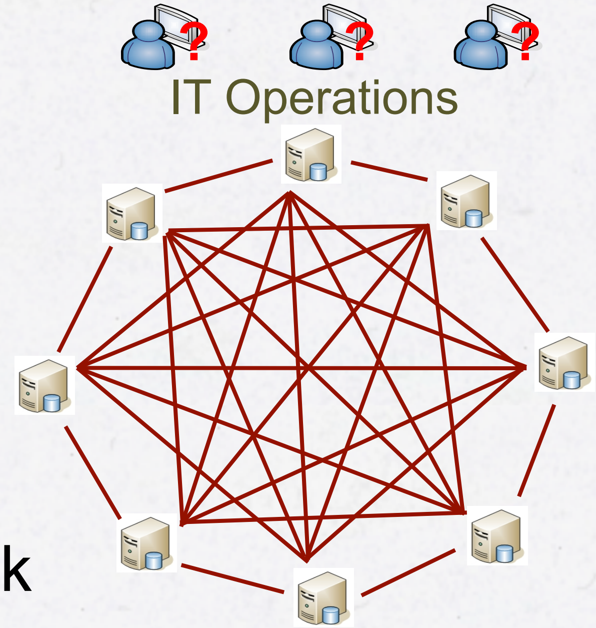
PROBLEM

- Many, IT organizations don’t have enough insight on what's happening in the black box.
- At the same time, as major IT breakdowns/incidents often is triggered by an unexpected combination of events that no one can really predict, or even thought of as a possible risk factor.
- The root cause analysis tends often to be time consuming...
- Difficult being proactive and analyzing trends...

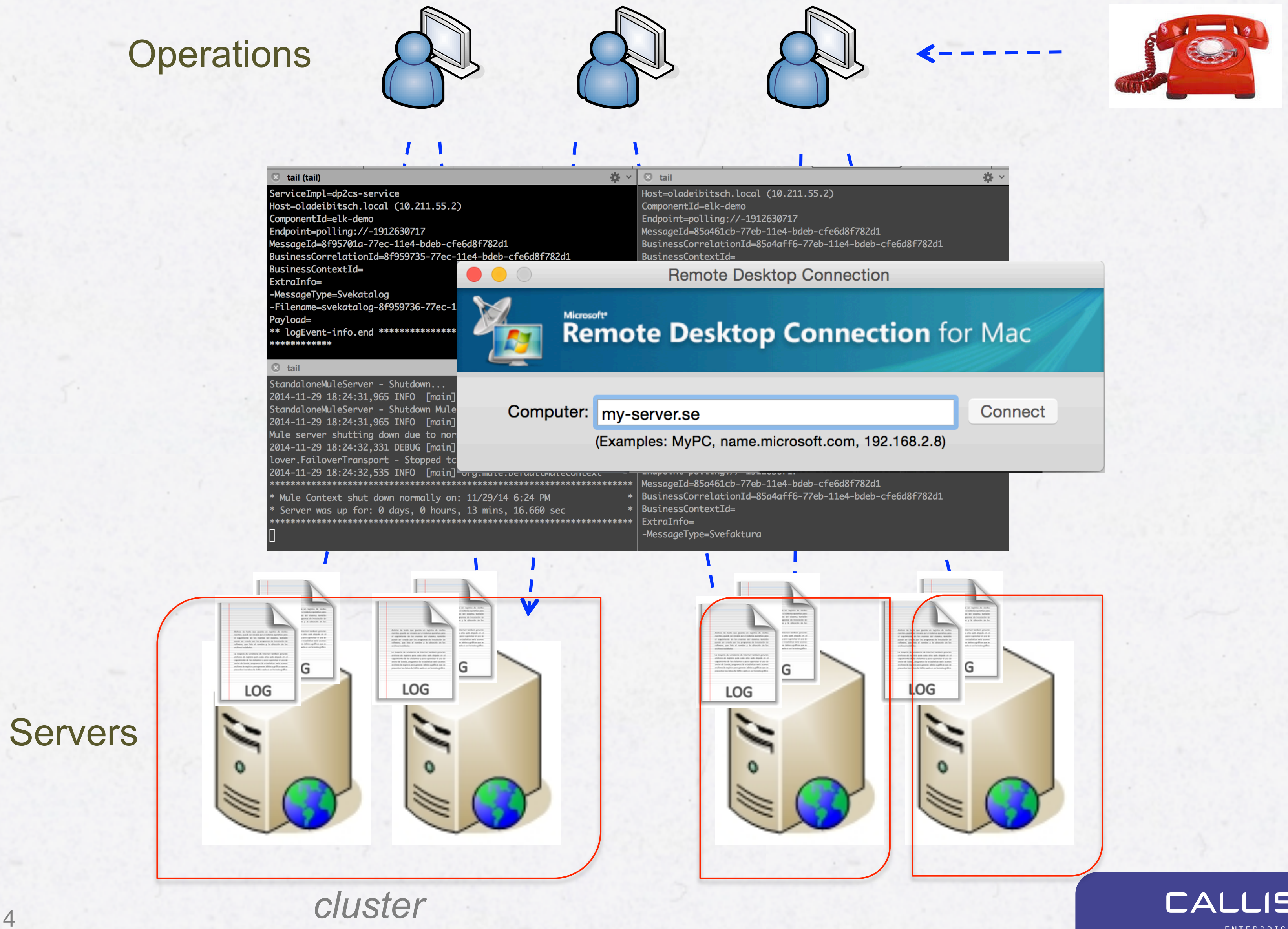


TECHNICAL CHALLANGES…

- Complex and distributed applications / servers
- Heterogeneous environments
- Restricted accessability
- Difficult to correlate events
- Troubleshooting – a needle in a haystack



ANALYSING THE ROOT CAUSE



tail (tail)

```
ServiceImpl=dp2cs-service
Host=oladeibitsch.local (10.211.55.2)
ComponentId=elk-demo
Endpoint=polling://-1912630717
MessageId=8f95701a-77ec-11e4-bdeb-cfe6d8f782d1
BusinessCorrelationId=8f959735-77ec-11e4-bdeb-cfe6d8f782d1
BusinessContextId=
ExtraInfo=
-MessageType=Svekatalog
-Filename=svekatalog-8f959736-77ec-1
Payload=
** logEvent-info.end **************
************
```

tail

```
StandaloneMuleServer - Shutdown...
2014-11-29 18:24:31,965 INFO  [main]
StandaloneMuleServer - Shutdown Mule
2014-11-29 18:24:31,965 INFO  [main]
Mule server shutting down due to nor
2014-11-29 18:24:32,331 DEBUG [main]
lover.FailoverTransport - Stopped tc
2014-11-29 18:24:32,535 INFO  [main] org.mule.DefaultMuleContext  -
**********************************************************************
* Mule Context shut down normally on: 11/29/14 6:24 PM               *
* Server was up for: 0 days, 0 hours, 13 mins, 16.660 sec            *
**********************************************************************
```

tail

```
Host=oladeibitsch.local (10.211.55.2)
ComponentId=elk-demo
Endpoint=polling://-1912630717
MessageId=85a461cb-77eb-11e4-bdeb-cfe6d8f782d1
BusinessCorrelationId=85a4aff6-77eb-11e4-bdeb-cfe6d8f782d1
BusinessContextId=
```

```
MessageId=85a461cb-77eb-11e4-bdeb-cfe6d8f782d1
BusinessCorrelationId=85a4aff6-77eb-11e4-bdeb-cfe6d8f782d1
BusinessContextId=
ExtraInfo=
-MessageType=Svefaktura
```

Remote Desktop Connection

Microsoft® Remote Desktop Connection for Mac

Computer: my-server.se

Connect

(Examples: MyPC, name.microsoft.com, 192.168.2.8)

LOG MANAGEMENT

LOG MANAGEMENT – LOG CHARACTERISTICS

```
2014-11-29 18:17:02,175 INFO  [dp2cs-service.stage1.02]
org.soitoolkit.commons.mule.messageLogger - soi-toolkit.log
** logEvent-info.start **********************************************************
IntegrationScenarioId=
ContractId=
LogMessage=msg-in
ServiceImpl=dp2cs-service
Host=oladeibitsch.local (10.211.55.2)
ComponentId=elk-demo
Endpoint=polling://-1912630717
MessageId=88a8a139-77eb-11e4-bdeb-cfe6d8f782d1
BusinessCorrelationId=88a8c854-77eb-11e4-bdeb-cfe6d8f782d1
BusinessContextId=
ExtraInfo=
-MessageType=Svekatalog
-Filename=svekatalog-88a8c855-77eb-11e4-bdeb-cfe6d8f782d1.txt
Payload=
** logEvent-info.end ************************************************************
```

LOG MANAGEMENT – LOG CHARACTERISTICS

```
127.0.0.1    - - [23/Nov/2014:06:42:29 +0100] "POST
  /vp/insuranceprocess//FindAllQuestions/1/rivtabp20 HTTP/1.1" 200 840 "-" "-"
127.0.0.1 - - [23/Nov/2014:06:42:29 +0100] "POST
  /vp/insuranceprocess//FindAllAnswers/1/rivtabp20 HTTP/1.1" 200 840 "-" "-"
127.0.0.1 - - [23/Nov/2014:06:42:29 +0100] "POST
  /vp/insuranceprocess//FindAllQuestions/1/rivtabp20 HTTP/1.1" 200 840 "-" "-”
```

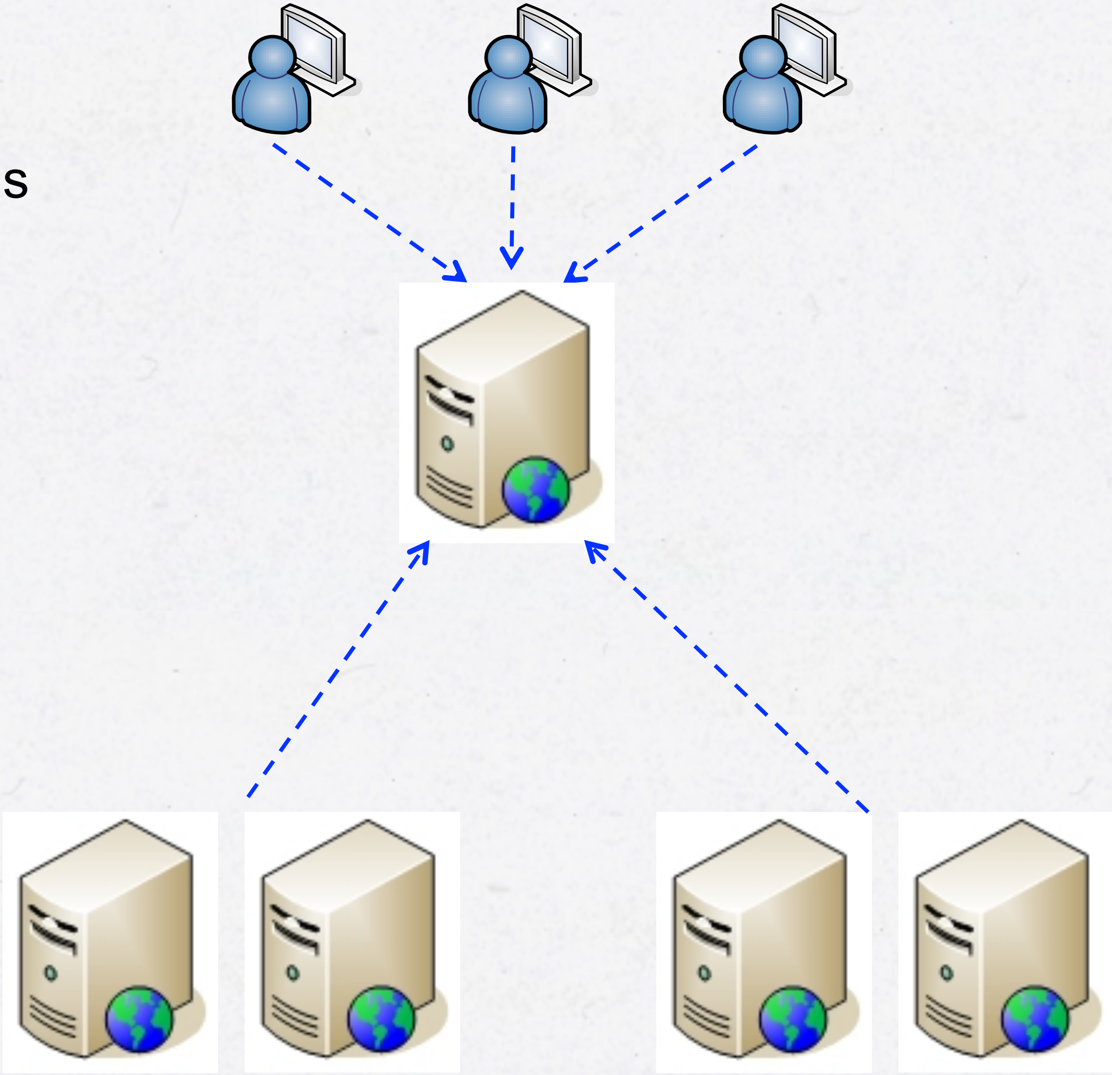
LOG MANAGEMENT – LOG CHARACTERISTICS

= TIMESTAMP + DATA

= ? + ?

REQUIREMENTS – LOG MONITORING

- Collecting Logs
- Parsing / Filter / Enrich Logs
- Centralize Logs
- Store Logs
- Analyze Logs
- Aggregate Logs
- Real-Time Analyse Logs
- Visualize Logs
- …



MEET ELASTICSEARCH, LOGSTASH AND KIBANA!

*”**Elasticsearch**, along with **Logstash** and **Kibana**, provides a powerful open source platform for indexing, searching and analyzing your data”*

Elasticsearch | Logstash | Kibana

MEET ELASTICSEARCH, LOGSTASH AND KIBANA!

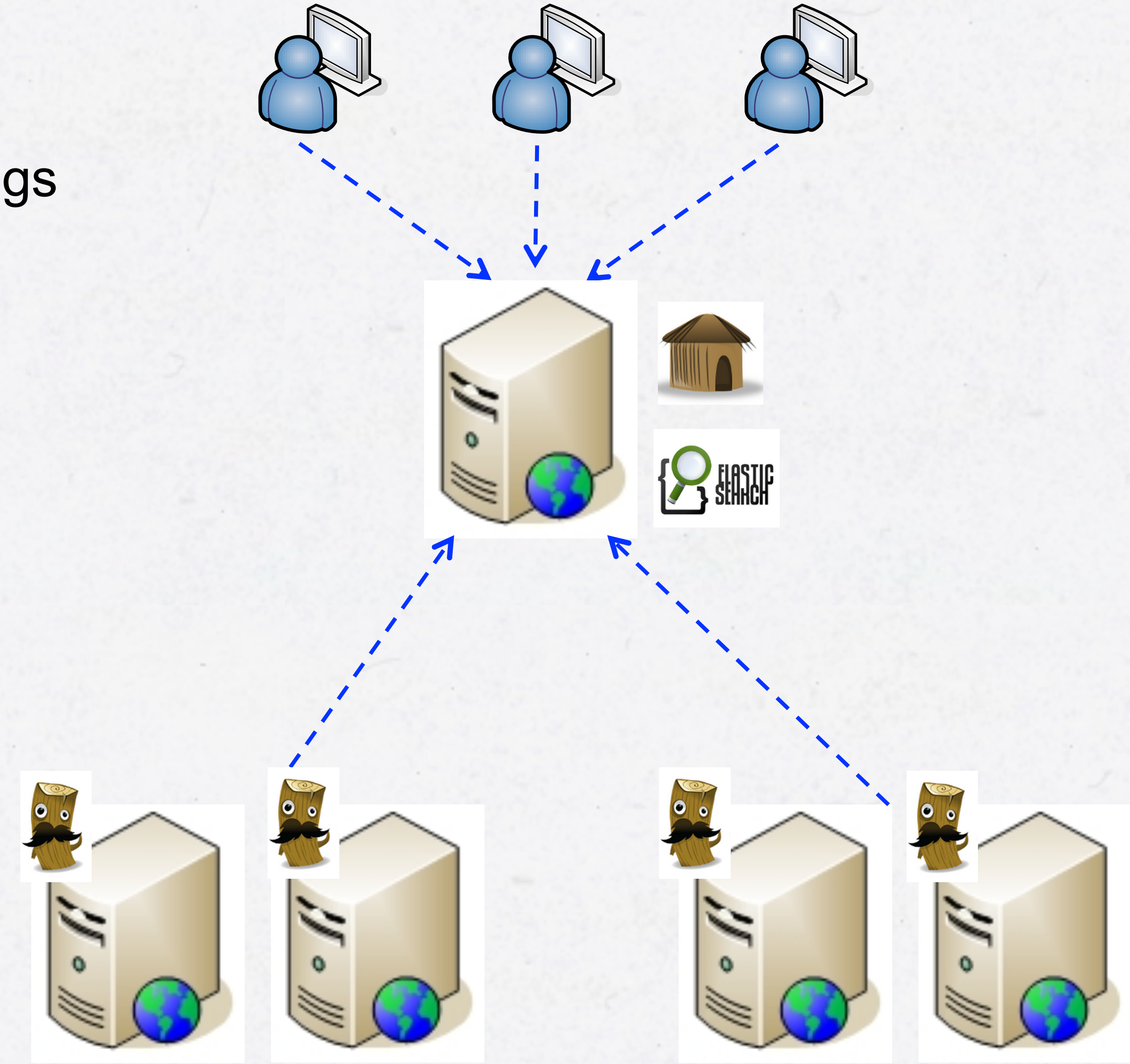
Elasticsearch: A document based search and analytics engine that makes data easy to explore using RESTful APIs.

Logstash: A event processing engine used for collecting, parsing and log enrichment.

Kibana: HTML 5 fronted, supporting dynamic dashboard(s), used to visualize Elasticsearch data.

REQUIREMENTS – LOG MONITORING

- ✓ Collecting Logs
- ✓ Parsing / Filter / Enrich Logs
- ✓ Centralize Logs
- ✓ Store Logs
- ✓ Analyze Logs
- ✓ Aggregate Logs
- ✓ Real-Time Analyse Logs
- ✓ Visualize Logs



ELASTICSEARCH

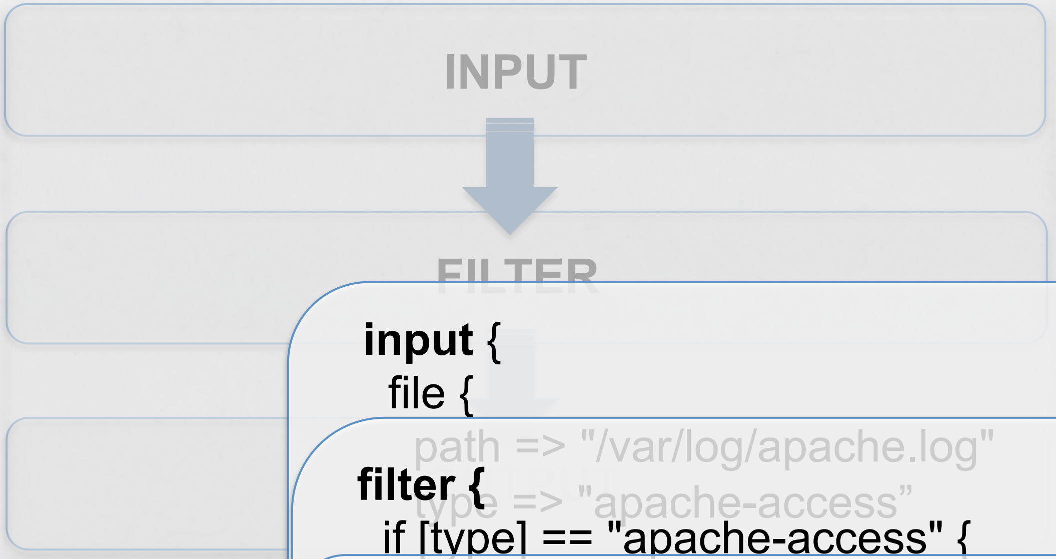
Elasticsearch is an open source RESTful search engine.

- Real time data
- Real time analytics
- High availability
- Scalability
- Document oriented
- RESTful API
- …

RESTful API
GET PUT POST DELETE
3891 \"
"agent" => "\"Mozilla/5.0 (Macintosh; Intel Mac OS X 10.9; rv:25.0) Gecko/20100101 Firefox/25.0\""

LOGSTASH - TERMINOLOGY

The logstash agent is a processing pipeline with three stages:



```
input {
  file {
    path => "/var/log/apache.log"
    type => "apache-access"
  }
  …
}
```

```
filter {
  if [type] == "apache-access" {
    grok {
      match => [ "message", "%{COMBINEDAPACHELOG}" ]
    }
  }
  …
}
```

```
output {
  elasticsearch { … }
  …
}
```

LOGSTASH – PLUGINS

<< INPUTS >>

collectd drupal_dblog elasticsearch eventlog exec ***file*** *ganglia gelf gemfire generator graphite heroku imap invalid_input irc jmx log4j lumberjack pipe puppet_facter rabbitmq rackspace redis relp s3 snmptrap sqlite sqs* ***stdin stomp*** *syslog tcp twitter udp unix varnishlog websocket wmi xmpp zenoss zeromq*

LOGSTASH – PLUGINS

<< FILTERS >>

advisor alter anonymize checksum cidr cipher clone collate csv **date** *dns* **drop** *elapsed elasticsearch environment extractnumbers fingerprint gelfify geoip grep* **grok** *grokdiscovery i18n json json_encode kv metaevent metrics* **multiline mutate** *noop prune punct railsparallelrequest range ruby sleep split sumnumbers syslog_pri throttle translate unique urldecode useragent uuid wms wmts xml zeromq*

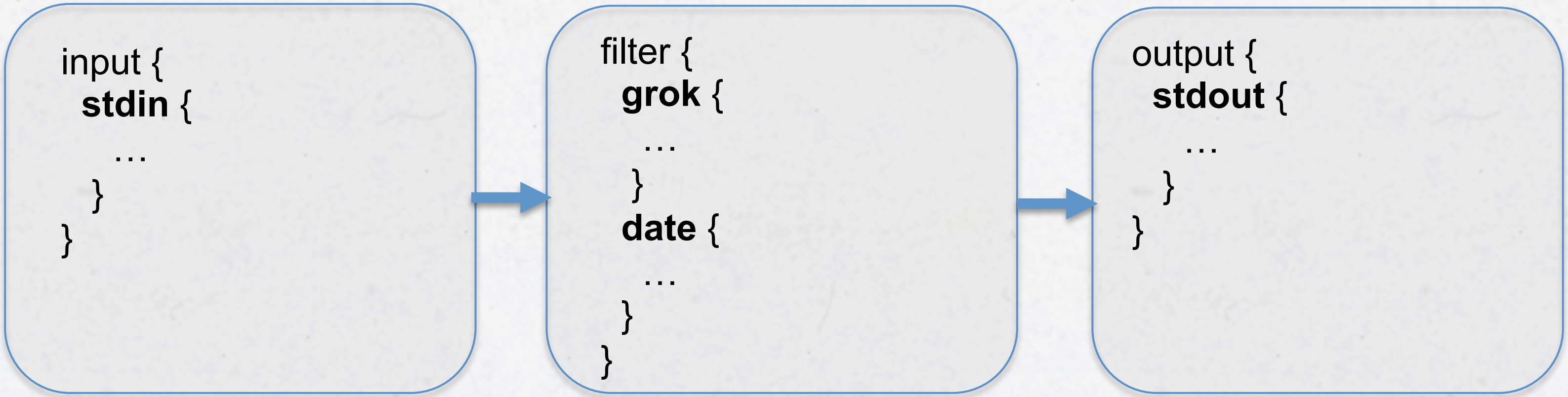
LOGSTASH – PLUGINS

<< OUTPUTS >>

boundary circonus cloudwatch csv datadog datadog_metrics ***elasticsearch*** *elasticsearch_http elasticsearch_river email exec file ganglia gelf gemfire google_bigquery google_cloud_storage graphite graphtastic hipchat http irc jira juggernaut librato loggly lumberjack metriccatcher mongodb nagios nagios_nsca null opentsdb pagerduty pipe rabbitmq rackspace redis redmine riak riemann s3 sns solr_http sqs statsd* ***stdout stomp*** *syslog tcp udp websocket xmpp zabbix zeromq*

DEMO 1 – A MINIMAL LOGSTASH CONFIGURATION

```
127.0.0.1 - - [23/Nov/2014:06:42:29 +0100] "POST
   /vp/insuranceprocess/FindAllQuestions/1/rivtabp20 HTTP/1.1" 200 840 "-" "-"
127.0.0.1 - - [23/Nov/2014:06:42:29 +0100] "POST
   /vp/insuranceprocess/FindAllAnswers/1/rivtabp20 HTTP/1.1" 200 840 "-" "-"
127.0.0.1 - - [23/Nov/2014:06:42:29 +0100] "POST
   /vp/insuranceprocess/FindAllQuestions/1/rivtabp20 HTTP/1.1" 200 840 "-" "-”
```



```
input {
  stdin {
     …
  }
}
```

```
filter {
  grok {
    …
  }
  date {
    …
  }
}
```

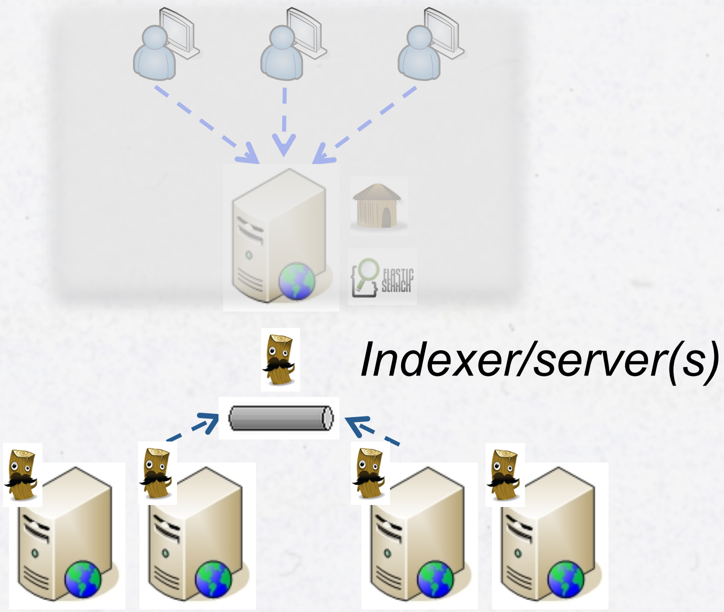
```
output {
  stdout {
    …
  }
}
```

DEMO 1 – A MINIMAL LOGSTASH CONFIGURATION (CONT.)

```
input {
    stdin {
        type => "apache-access"
    }
}
filter {
    if [type] == "apache-access" {
        grok {
            match => { "message" => "%{COMBINEDAPACHELOG}" }
        }
        date {
            match => [ "timestamp" , "dd/MMM/yyyy:HH:mm:ss Z" ]
        }
    }
}
output {
    stdout { codec => rubydebug }
}
```

LOGSTASH – ARCHITECTUAL OVERVIEW

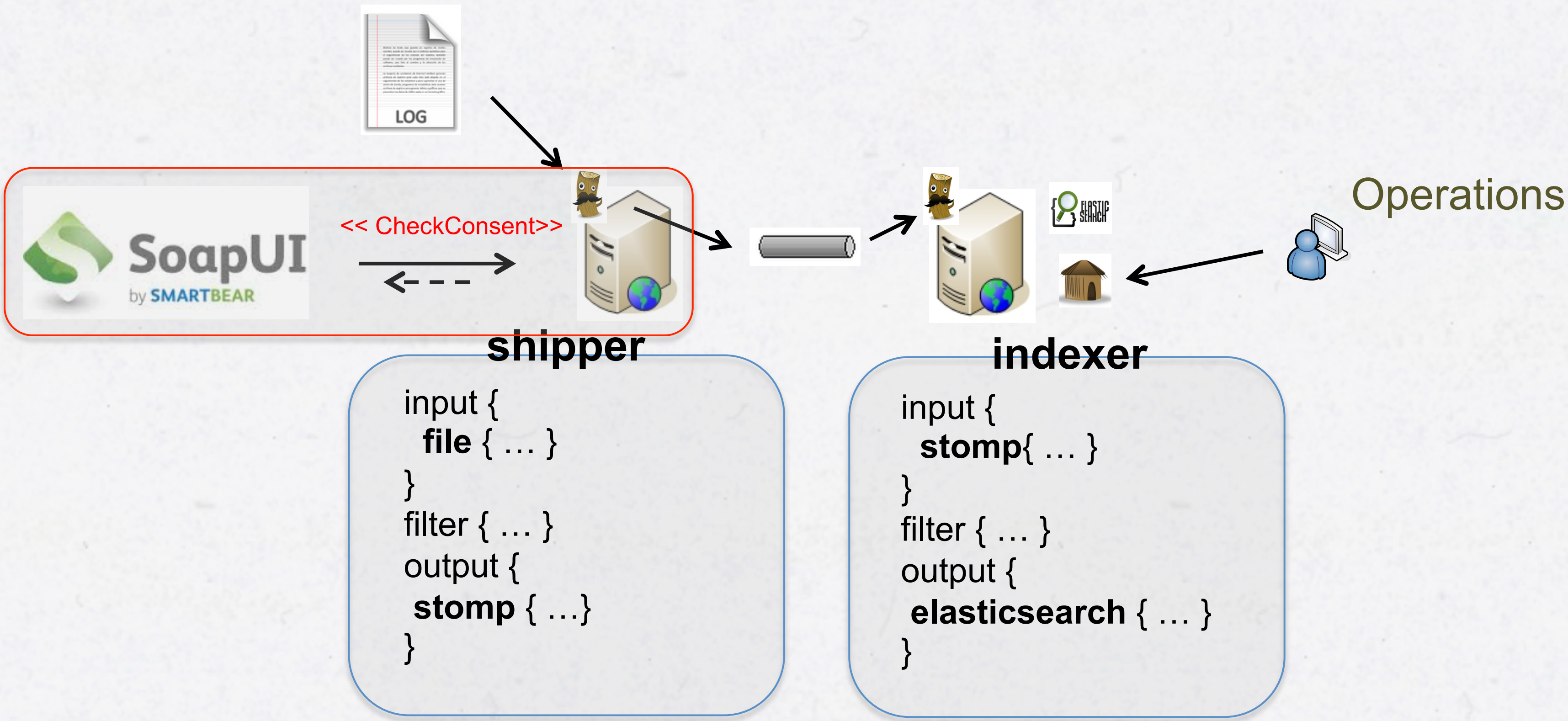
- ”Shipper/agents”
  - Ships logs to logstash server, logstash remote agents
- ”Indexer/server”
  - Receives and indexes the events within logstash server.
  - Logstash servers run one or more of the components independently, which helps to separate components and scale logstash



DEMO 2 – VISUALIZING WITH KIBANA

The purpose of this demo is to show how to start visualizing logs in Kibana using panels like:

- Tables, Histograms, Terms

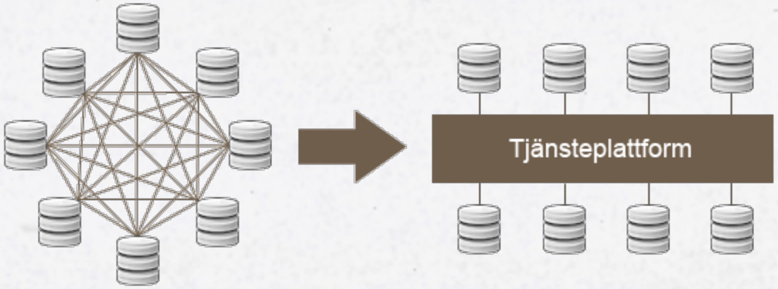


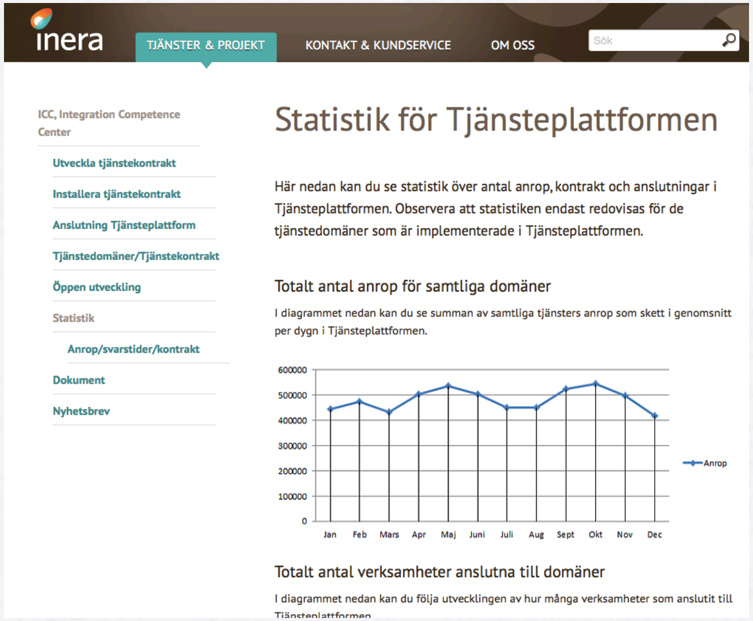
```
input {
  file { … }
}
filter { … }
output {
 stomp { …}
}
```

```
input {
  stomp{ … }
}
filter { … }
output {
 elasticsearch { … }
}
```

CASE STUDY – SKLTP

SKLTP is an open source project that implements priority parts of a service platform according to the reference architecture for health care. SKLTP used by Inera in the national service platform. SKLTP is also used as a regional service platform in different counties.




inera
TJÄNSTER & PROJEKT
KONTAKT & KUNDSERVICE
OM OSS
Sök
ICC, Integration Competence Center
Utveckla tjänstekontrakt
Installera tjänstekontrakt
Anslutning Tjänsteplattform
Tjänstedomäner/Tjänstekontrakt
Öppen utveckling
Statistik
Anrop/svarstider/kontrakt
Dokument
Nyhetsbrev

Statistik för Tjänsteplattformen

Här nedan kan du se statistik över antal anrop, kontrakt och anslutningar i Tjänsteplattformen. Observera att statistiken endast redovisas för de tjänstedomäner som är implementerade i Tjänsteplattformen.

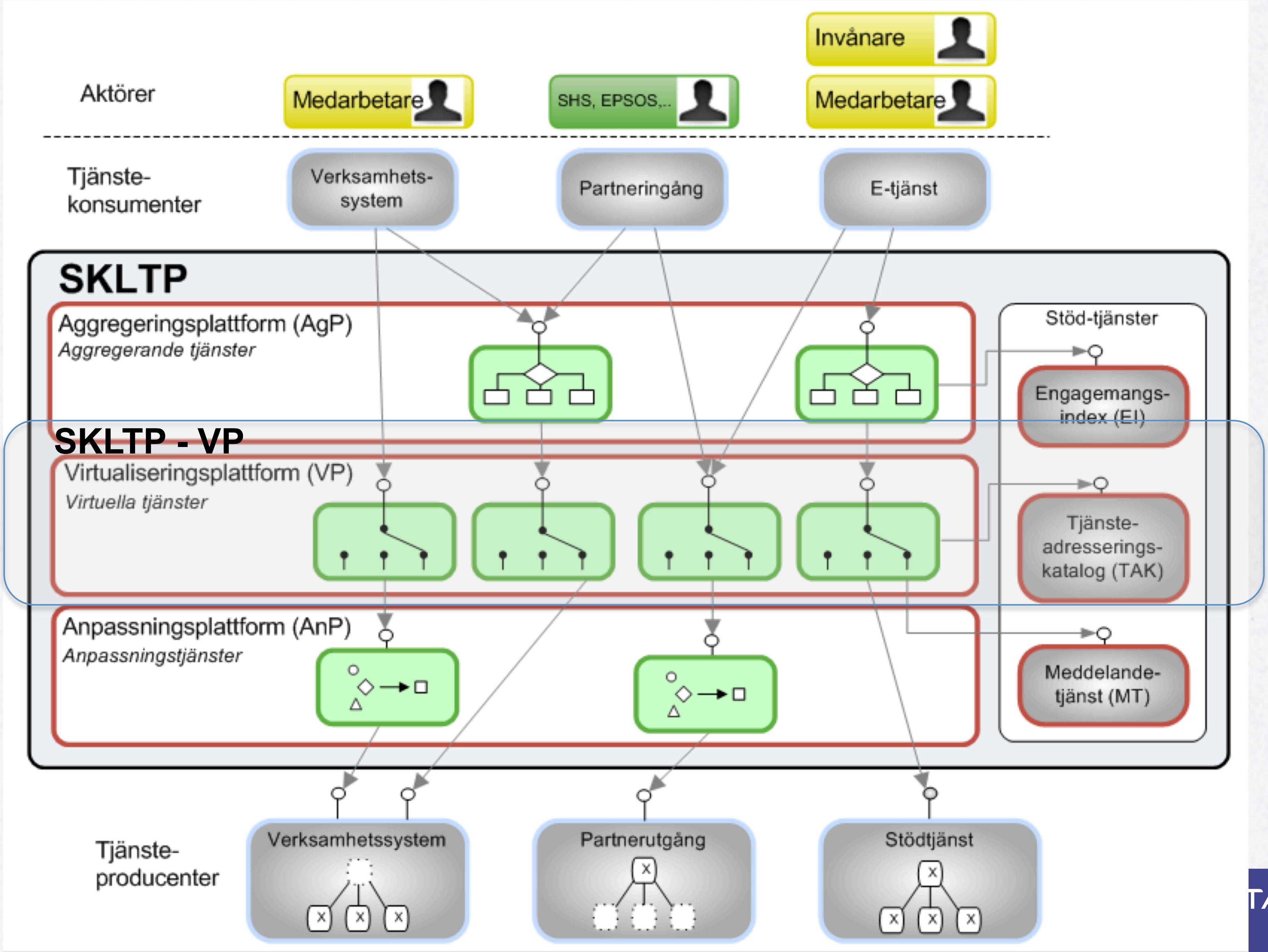
Totalt antal anrop för samtliga domäner

I diagrammet nedan kan du se summan av samtliga tjänsters anrop som skett i genomsnitt per dygn i Tjänsteplattformen.

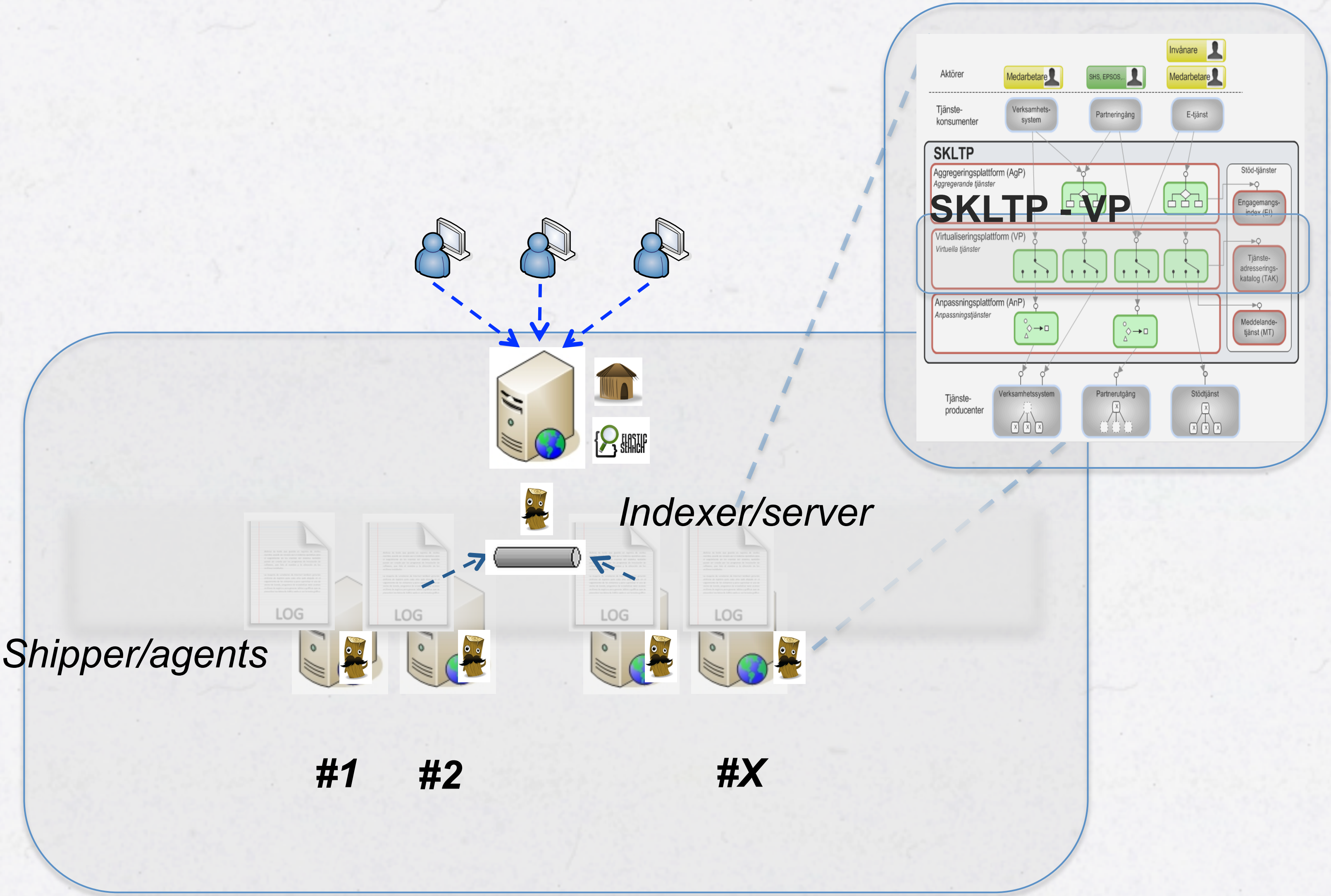
Totalt antal verksamheter anslutna till domäner

I diagrammet nedan kan du följa utvecklingen av hur många verksamheter som anslutit till Tjänsteplattformen

CASE STUDY – SKLTP (CONT.)



CASE STUDY – MONITORING ”THE BLACK BOX”



SKLTP - VP

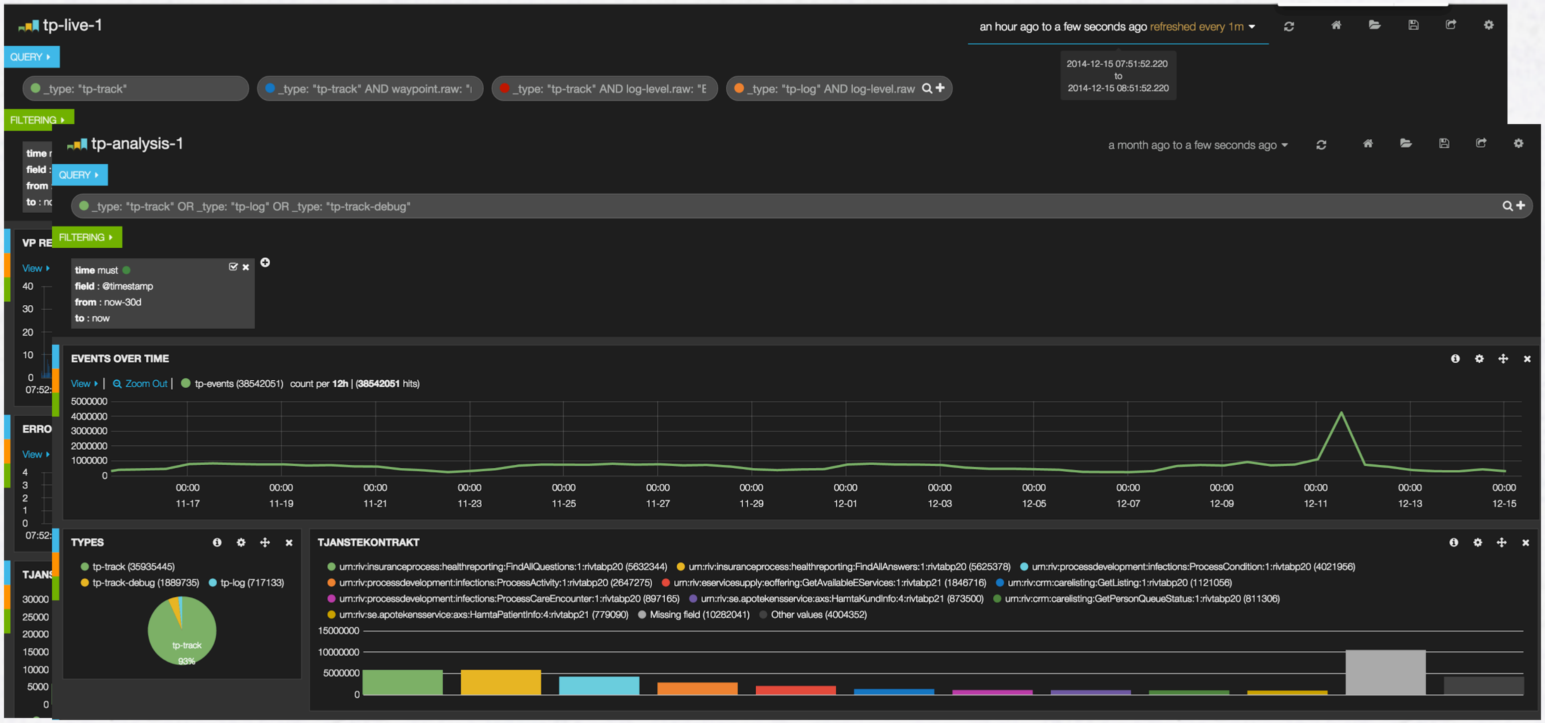
Indexer/server

Shipper/agents

#1 #2 #X

DEMO 3 – SKLTP MONITORING

”real time” is the only time…



tp-live-1

an hour ago to a few seconds ago refreshed every 1m

2014-12-15 07:51:52.220
to
2014-12-15 08:51:52.220

QUERY

- ● _type: "tp-track"
- ● _type: "tp-track" AND waypoint.raw: "
- ● _type: "tp-track" AND log-level.raw: "E
- ● _type: "tp-log" AND log-level.raw

tp-analysis-1

a month ago to a few seconds ago

QUERY

● _type: "tp-track" OR _type: "tp-log" OR _type: "tp-track-debug"

FILTERING

time must ●
field : @timestamp
from : now-30d
to : now

EVENTS OVER TIME

● tp-events (38542051) count per **12h** | (**38542051** hits)

TYPES

- ● tp-track (35935445)
- ● tp-track-debug (1889735)
- ● tp-log (717133)

TJANSTEKONTRAKT

- ● urn:riv:insuranceprocess:healthreporting:FindAllQuestions:1:rivtabp20 (5632344)
- ● urn:riv:insuranceprocess:healthreporting:FindAllAnswers:1:rivtabp20 (5625378)
- ● urn:riv:processdevelopment:infections:ProcessCondition:1:rivtabp20 (4021956)
- ● urn:riv:processdevelopment:infections:ProcessActivity:1:rivtabp20 (2647275)
- ● urn:riv:eservicesupply:eoffering:GetAvailableEServices:1:rivtabp21 (1846716)
- ● urn:riv:crm:carelisting:GetListing:1:rivtabp20 (1121056)
- ● urn:riv:processdevelopment:infections:ProcessCareEncounter:1:rivtabp20 (897165)
- ● urn:riv:se.apotekensservice:axs:HamtaKundInfo:4:rivtabp21 (873500)
- ● urn:riv:crm:carelisting:GetPersonQueueStatus:1:rivtabp20 (811306)
- ● urn:riv:se.apotekensservice:axs:HamtaPatientInfo:4:rivtabp21 (779090)
- ● Missing field (10282041)
- ● Other values (4004352)

analysing events over time…

SUMMARY

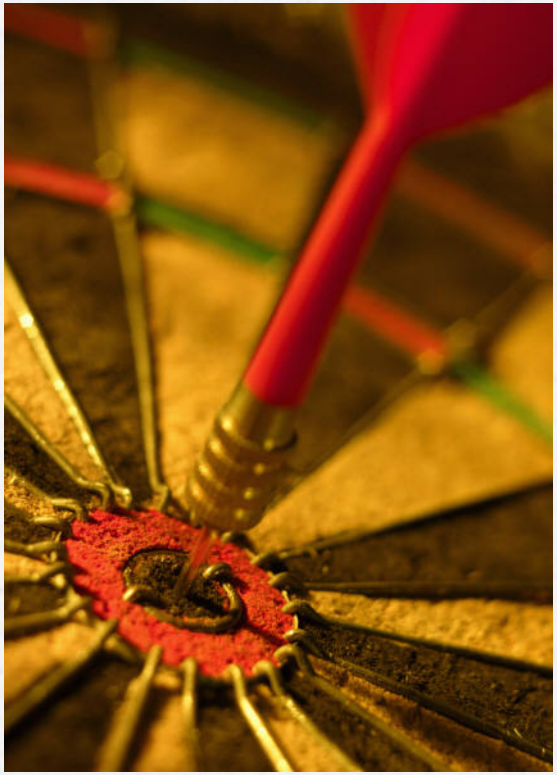
The **ELK** stack is three seamlessly integrated open source products…

…that helps us to ***centralize***, ***consolidate, structure*** and ***visualize*** logs…

…which enables us to:

- ✓perform troubleshooting
- ✓perform log analysis
- ✓work proactively

➔ LOG DATA IS UNUSED, USE IT!



CALLISTA
— ENTERPRISE —