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 CHALLENGES…

- Complex and distributed applications / servers
- Heterogeneous environments
- Restricted accessibility
- Difficult to correlate events
- Troubleshooting – a needle in a haystack
ANALYSING THE ROOT CAUSE
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 *******************************************************
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

\[ \text{LOG} = \text{TIMESTAMP} + \text{DATA} \]
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"
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 is an open source RESTful search engine.

- Real time data
- Real time analytics
- High availability
- Scalability
- Document oriented
- RESTful API
- ...
The logstash agent is a processing pipeline with three stages:

```
in
  input {
    file {
      path => "/var/log/apache.log"
    }
  } fi
   
  filter {
    if [type] == "apache-access"
      grok {
        match => 
    }}

  output {
    elasticsearch {
      path => [ "message", "%{COMBINEDAPACHELOG}" ]
    }
  }
```
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

```
input {
  stdin {
    ...
  }
}

filter {
  grok {
    ...
  }
  date {
    ...
  }
}

output {
  stdout {
    ...
  }
}
```

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 {
        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 – ARCHITECTURAL 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
The purpose of this demo is to show how to start visualizing logs in Kibana using panels like:
- Tables, Histograms, Terms
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.
CASE STUDY – SKLTP (CONT.)
CASE STUDY – MONITORING "THE BLACK BOX"
"real time" is the only time…

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!**